GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: September 9, 2004, 22:21:53 ; Search time 63 Seconds
(without alignments)
769.991 Million cell updates/sec

Title: US-09-887-855-2
Perfect score: 2000
Sequence: 1 MRPGTALQAVLLAVLLVGLR.....PDQMGRSKESGWVENEIYGY 374

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 704768 seqs, 129704438 residues

Total number of hits satisfying chosen parameters: 704768

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending_Patents_AA_New:*
1: /cgn2_6/ptodata/2/paa/PCT_NEW_COMB.pep:*
2: /cgn2_6/ptodata/2/paa/US06_NEW_COMB.pep:*
3: /cgn2_6/ptodata/2/paa/US07_NEW_COMB.pep:*
4: /cgn2_6/ptodata/2/paa/US08_NEW_COMB.pep:*
5: /cgn2_6/ptodata/2/paa/US09_NEW_COMB.pep:*
6: /cgn2_6/ptodata/2/paa/US10_NEW_COMB.pep:*
7: /cgn2_6/ptodata/2/paa/US60_NEW_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1990	99.5	374	6	US-10-100-683-8305	Sequence 8305, Ap
2	1986	99.3	382	6	US-10-797-366-137	Sequence 137, App
3	1986	99.3	382	6	US-10-771-187-137	Sequence 137, App
4	1982	99.1	387	1	PCT-US04-09202-861	Sequence 861, App
5	1979	99.0	374	1	PCT-US04-09202-1040	Sequence 1040, Ap
6	1805	90.2	356	1	PCT-US04-09202-363	Sequence 363, App
7	1784	89.2	358	6	US-10-170-205E-35136	Sequence 35136, A
8	1765.5	88.3	389	1	PCT-US04-09202-1010	Sequence 1010, Ap
9	1198.5	59.9	260	6	US-10-897-911-3	Sequence 3, Appli
10	1198.5	59.9	260	6	US-10-898-615-3	Sequence 3, Appli
11	645	32.2	273	6	US-10-152-372-540	Sequence 540, App
12	645	32.2	273	6	US-10-123-155-540	Sequence 540, App
13	645	32.2	273	6	US-10-063-685-168	Sequence 168, App
14	645	32.2	273	6	US-10-143-117-540	Sequence 540, App
15	645	32.2	273	6	US-10-127-823A-540	Sequence 540, App
16	645	32.2	273	6	US-10-884-091-18	Sequence 18, Appl
17	645	32.2	273	6	US-10-897-911-2	Sequence 2, Appli
18	645	32.2	273	6	US-10-140-923-540	Sequence 540, App
19	645	32.2	273	6	US-10-898-615-2	Sequence 2, Appli
20	645	32.2	273	6	US-10-142-426-540	Sequence 540, App
21	645	32.2	273	6	US-10-141-756-540	Sequence 540, App
22	645	32.2	273	6	US-10-919-654-3	Sequence 3, Appli
23	645	32.2	273	6	US-10-919-654-12	Sequence 12, Appl
24	645	32.2	273	6	US-10-919-654-15	Sequence 15, Appl
25	645	32.2	273	6	US-10-919-654-56	Sequence 56, Appl
26	645	32.2	273	6	US-10-141-761-540	Sequence 540, App
27	645	32.2	273	6	US-10-140-472-540	Sequence 540, App
28	645	32.2	273	6	US-10-140-805-540	Sequence 540, App
29	645	32.2	273	6	US-10-140-864-540	Sequence 540, App
30	645	32.2	273	6	US-10-141-759-540	Sequence 540, App
31	645	32.2	273	6	US-10-141-761-540	Sequence 540, App
32	645	32.2	273	6	US-10-142-885-540	Sequence 540, App
33	645	32.2	273	6	US-10-146-731-540	Sequence 540, App
34	576	28.8	232	6	US-10-170-205E-32318	Sequence 32318, A
35	576	28.8	232	6	US-10-917-503-14520	Sequence 14520, A
36	576	28.8	232	6	US-10-919-654-5	Sequence 5, Appli
37	576	28.8	232	6	US-10-919-654-7	Sequence 7, Appli
38	576	28.8	232	6	US-10-919-654-13	Sequence 13, Appl
39	576	28.8	232	6	US-10-919-654-16	Sequence 16, Appl
40	576	28.8	232	6	US-10-919-654-63	Sequence 63, Appl
41	576	28.8	232	6	US-10-919-654-64	Sequence 64, Appl
42	576	28.8	232	6	US-10-919-654-65	Sequence 65, Appl
43	576	28.8	232	6	US-10-919-654-69	Sequence 69, Appl
44	576	28.8	232	6	US-10-919-654-70	Sequence 70, Appl
45	576	28.8	232	6	US-10-919-654-71	Sequence 71, Appl

ALIGNMENTS

RESULT 1
US-10-100-683-8305
; Sequence 8305, Application US/10100683
; GENERAL INFORMATION:
; APPLICANT: Rosen, et al.
; TITLE OF INVENTION: Human Secreted Proteins
; FILE REFERENCE: PS900
; CURRENT APPLICATION NUMBER: US/10/100,683
; CURRENT FILING DATE: 2002-03-19
; PRIOR APPLICATION NUMBER: US 60/040,162
; PRIOR FILING DATE: 1997-03-07
; PRIOR APPLICATION NUMBER: US 60/043,576
; PRIOR FILING DATE: 1997-04-11
; PRIOR APPLICATION NUMBER: US 60/047,601
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: US 60/056,845
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: US 60/043,580
; PRIOR FILING DATE: 1997-04-11
; PRIOR APPLICATION NUMBER: US 60/047,599
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: US 60/056,664
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: US 60/043,314
; PRIOR FILING DATE: 1997-04-11
; PRIOR APPLICATION NUMBER: US 60/047,632
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: US 60/056,892
; PRIOR FILING DATE: 1997-08-22
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 13468
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8305
; LENGTH: 374
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (84)
; OTHER INFORMATION: Xaa equals any amino acid
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (112)
; OTHER INFORMATION: Xaa equals any amino acid
US-10-100-683-8305

Query Match 99.5%; Score 1990; DB 6; Length 374;
Best Local Similarity 99.5%; Pred. No. 1.2e-171;
Matches 372; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy          1 MRPGTALQAVLLAVLLVGLRAATGRLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAK 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MRPGTALQAVLLAVLLVGLRAATGRLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAK 60

Qy         61 EACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWT 120
              |||||||||||||||||||||| |||||||||||||||||||||||||||| ||||||||
Db         61 EACRRDGGQLVSIESEDEQKLIEXFIENLLPSDGDFWIGLRRREEKQSNSTXCQDLYAWT 120

Qy        121 DGSISQFRNWYVDEPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSD 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 DGSISQFRNWYVDEPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSD 180

Qy        181 EKPAVPSREAEGEETELTTPVLPEETQEEDAKKTFKESREAALNLAYILIPSIPLLLLLV 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 EKPAVPSREAEGEETELTTPVLPEETQEEDAKKTFKESREAALNLAYILIPSIPLLLLLV 240

Qy        241 VTTVVCWVWICRKRKREQPDPSTKKQHTIWPSPHQGNSPDLEVYNVIRKQSEADLAETRP 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 VTTVVCWVWICRKRKREQPDPSTKKQHTIWPSPHQGNSPDLEVYNVIRKQSEADLAETRP 300

Qy        301 DLKNISFRVCSGEATPDDMSCDYDNMAVNPSESGFVTLVSVESGFVTNDIYEFSPDQMGR 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 DLKNISFRVCSGEATPDDMSCDYDNMAVNPSESGFVTLVSVESGFVTNDIYEFSPDQMGR 360

Qy        361 SKESGWVENEIYGY 374
              ||||||||||||||
Db        361 SKESGWVENEIYGY 374
```

```
RESULT 2
US-10-797-366-137
; Sequence 137, Application US/10797366
; GENERAL INFORMATION:
;  APPLICANT: Genentech, Inc.
;  APPLICANT:  Ashkenazi, Avi
;  APPLICANT:  Botstein, David
;  APPLICANT:  Desnoyers, Luc
;  APPLICANT:  Eaton, Dan L.
;  APPLICANT:  Ferrara, Napoleone
;  APPLICANT:  Filvaroff, Ellen
;  APPLICANT:  Fong, Sherman
;  APPLICANT:  Gao, Wei-Qiang
;  APPLICANT:  Gerber, Hanspeter
;  APPLICANT:  Gerritsen, Mary E.
;  APPLICANT:  Goddard, A.
;  APPLICANT:  Godowski, Paul J.
;  APPLICANT:  Grimaldi, Christopher J.
;  APPLICANT:  Gurney, Austin L.
;  APPLICANT:  Hillan, Kenneth, J.
;  APPLICANT:  Kljavin, Ivar J.
;  APPLICANT:  Mather, Jennie P.
;  APPLICANT:  Pan, James
;  APPLICANT:  Paoni, Nicholas F.
;  APPLICANT:  Roy, Margaret Ann
;  APPLICANT:  Stewart, Timothy A.
;  APPLICANT:  Tumas, Daniel
;  APPLICANT:  Williams, P. Mickey
;  APPLICANT:  Wood, William, I.
;  TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
;  TITLE OF INVENTION:  Acids Encoding the Same
;  FILE REFERENCE: 10466-14
;  CURRENT APPLICATION NUMBER: US/10/797,366
;  CURRENT FILING DATE:  2004-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/04414
;  PRIOR FILING DATE: 2000-02-22
;  PRIOR APPLICATION NUMBER: US 60/143,048
;  PRIOR FILING DATE: 1999-07-07
;  PRIOR APPLICATION NUMBER: US 60/145,698
;  PRIOR FILING DATE: 1999-07-26
;  PRIOR APPLICATION NUMBER: US 60/146,222
;  PRIOR FILING DATE: 1999-07-28
;  PRIOR APPLICATION NUMBER: PCT/US99/20594
;  PRIOR FILING DATE: 1999-09-08
;  PRIOR APPLICATION NUMBER: PCT/US99/20944
;  PRIOR FILING DATE: 1999-09-13
;  PRIOR APPLICATION NUMBER: PCT/US99/21090
;  PRIOR FILING DATE: 1999-09-15
;  PRIOR APPLICATION NUMBER: PCT/US99/21547
;  PRIOR FILING DATE: 1999-09-15
;  PRIOR APPLICATION NUMBER: PCT/US99/23089
;  PRIOR FILING DATE: 1999-10-05
;  PRIOR APPLICATION NUMBER: PCT/US99/28214
;  PRIOR FILING DATE: 1999-11-29
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 137
;   LENGTH: 382
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-797-366-137

  Query Match             99.3%;  Score 1986;  DB 6;  Length 382;
  Best Local Similarity   97.9%;  Pred. No. 2.9e-171;
  Matches  374;  Conservative    0;  Mismatches    0;  Indels    8;  Gaps    1;
```

```
Qy          1 MRPGTALQAVLLAVLLVGLRAATGRLLS--------GQPVCRGGTQRPCYKVIYFHDTSR 52
              ||||||||||||||||||||||||||||        ||||||||||||||||||||||||
Db          1 MRPGTALQAVLLAVLLVGLRAATGRLLSASDLDLRGGQPVCRGGTQRPCYKVIYFHDTSR 60

Qy         53 RLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTA 112
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 RLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTA 120

Qy        113 CQDLYAWTDGSISQFRNWYVDEPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKN 172
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CQDLYAWTDGSISQFRNWYVDEPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKN 180

Qy        173 NFICKYSDEKPAVPSREAEGEETELTTPVLPEETQEEDAKKTFKESREAALNLAYILIPS 232
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 NFICKYSDEKPAVPSREAEGEETELTTPVLPEETQEEDAKKTFKESREAALNLAYILIPS 240

Qy        233 IPLLLLLVVTTVVCWVWICRKRKREQPDPSTKKQHTIWPSPHQGNSPDLEVYNVIRKQSE 292
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 IPLLLLLVVTTVVCWVWICRKRKREQPDPSTKKQHTIWPSPHQGNSPDLEVYNVIRKQSE 300

Qy        293 ADLAETRPDLKNISFRVCSGEATPDDMSCDYDNMAVNPSESGFVTLVSVESGFVTNDIYE 352
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ADLAETRPDLKNISFRVCSGEATPDDMSCDYDNMAVNPSESGFVTLVSVESGFVTNDIYE 360

Qy        353 FSPDQMGRSKESGWVENEIYGY 374
              ||||||||||||||||||||||
Db        361 FSPDQMGRSKESGWVENEIYGY 382
```

```
RESULT 3
US-10-771-187-137
; Sequence 137, Application US/10771187
; GENERAL INFORMATION:
;  APPLICANT: Genentech, Inc.
;  APPLICANT:  Ashkenazi, Avi
;  APPLICANT:  Botstein, David
;  APPLICANT:  Desnoyers, Luc
;  APPLICANT:  Eaton, Dan L.
;  APPLICANT:  Ferrara, Napoleone
;  APPLICANT:  Filvaroff, Ellen
;  APPLICANT:  Fong, Sherman
;  APPLICANT:  Gao, Wei-Qiang
;  APPLICANT:  Gerber, Hanspeter
;  APPLICANT:  Gerritsen, Mary E.
;  APPLICANT:  Goddard, A.
;  APPLICANT:  Godowski, Paul J.
;  APPLICANT:  Grimaldi, Christopher J.
;  APPLICANT:  Gurney, Austin L.
;  APPLICANT:  Hillan, Kenneth, J.
;  APPLICANT:  Kljavin, Ivar J.
```

```
;  APPLICANT:  Mather, Jennie P.
;  APPLICANT:  Pan, James
;  APPLICANT:  Paoni, Nicholas F.
;  APPLICANT:  Roy, Margaret Ann
;  APPLICANT:  Stewart, Timothy A.
;  APPLICANT:  Tumas, Daniel
;  APPLICANT:  Williams, P. Mickey
;  APPLICANT:  Wood, William, I.
;  TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
;  TITLE OF INVENTION:  Acids Encoding the Same
;  FILE REFERENCE: 39780-1618P2C78C1
;  CURRENT APPLICATION NUMBER: US/10/771,187
;  CURRENT FILING DATE:  2004-02-02
;  PRIOR APPLICATION NUMBER: 09/909,064
;  PRIOR FILING DATE: 2001-07-18
;  PRIOR APPLICATION NUMBER: 09/665,350
;  PRIOR FILING DATE: 2000-09-18
;  PRIOR APPLICATION NUMBER: PCT/US00/04414
;  PRIOR FILING DATE: 2000-02-22
;  PRIOR APPLICATION NUMBER: PCT/US98/19437
;  PRIOR FILING DATE: 1998-09-17
;  PRIOR APPLICATION NUMBER: PCT/US98/19330
;  PRIOR FILING DATE: 1998-09-16
;  PRIOR APPLICATION NUMBER: 60/088,026
;  PRIOR FILING DATE: 1998-06-04
;  PRIOR APPLICATION NUMBER: 60/066,770
;  PRIOR FILING DATE: 1997-11-24
;  PRIOR APPLICATION NUMBER: 60/065,186
;  PRIOR FILING DATE: 1997-11-12
;  NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 137
;   LENGTH: 382
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-771-187-137

  Query Match           99.3%;  Score 1986;  DB 6;  Length 382;
  Best Local Similarity  97.9%;  Pred. No. 2.9e-171;
  Matches  374;  Conservative    0;  Mismatches    0;  Indels    8;  Gaps    1;

Qy          1 MRPGTALQAVLLAVLLVGLRAATGRLLS--------GQPVCRGGTQRPCYKVIYFHDTSR 52
              ||||||||||||||||||||||||||||        ||||||||||||||||||||||||
Db          1 MRPGTALQAVLLAVLLVGLRAATGRLLSASDLDLRGGQPVCRGGTQRPCYKVIYFHDTSR 60

Qy         53 RLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTA 112
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 RLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTA 120

Qy        113 CQDLYAWTDGSISQFRNWYVDEPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKN 172
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CQDLYAWTDGSISQFRNWYVDEPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKN 180

Qy        173 NFICKYSDEKPAVPSREAEGEETELTTPVLPEETQEEDAKKTFKESREAALNLAYILIPS 232
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 NFICKYSDEKPAVPSREAEGEETELTTPVLPEETQEEDAKKTFKESREAALNLAYILIPS 240

Qy        233 IPLLLLLVVTTVVCWVWICRKRKREQPDPSTKKQHTIWPSPHQGNSPDLEVYNVIRKQSE 292
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 IPLLLLLVVTTVVCWVWICRKRKREQPDPSTKKQHTIWPSPHQGNSPDLEVYNVIRKQSE 300

Qy        293 ADLAETRPDLKNISFRVCSGEATPDDMSCDYDNMAVNPSESGFVTLVSVESGFVTNDIYE 352
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ADLAETRPDLKNISFRVCSGEATPDDMSCDYDNMAVNPSESGFVTLVSVESGFVTNDIYE 360

Qy        353 FSPDQMGRSKESGWVENEIYGY 374
              ||||||||||||||||||||||
Db        361 FSPDQMGRSKESGWVENEIYGY 382


RESULT 4
PCT-US04-09202-861
; Sequence 861, Application PC/TUS0409202
; GENERAL INFORMATION:
;  APPLICANT: Zhou, Ping
;  APPLICANT:  Tang, Y. Tom
;  APPLICANT:  Hu, Tianhua
;  APPLICANT:  Wang, Jian-Rui
;  APPLICANT:  Wang, Zhi Wei
;  APPLICANT:  Drmanac, Radoje T.
;  TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
;  FILE REFERENCE: 824CIP/PCT
;  CURRENT APPLICATION NUMBER: PCT/US04/09202
;  CURRENT FILING DATE:  2004-04-06
;  PRIOR APPLICATION NUMBER: US 60/458,824
;  PRIOR FILING DATE: 2003-03-28
;  PRIOR APPLICATION NUMBER: US 10/296,115
;  PRIOR FILING DATE: 2000-12-22
;  PRIOR APPLICATION NUMBER: US 10/275,027
;  PRIOR FILING DATE: 2001-01-25
;  PRIOR APPLICATION NUMBER: US 10/276,774
;  PRIOR FILING DATE: 2001-02-05
;  PRIOR APPLICATION NUMBER: US 10/220,366
;  PRIOR FILING DATE: 2001-02-26
;  PRIOR APPLICATION NUMBER: US 10/221,279
;  PRIOR FILING DATE: 2001-03-05
;  PRIOR APPLICATION NUMBER: US 10/450,763
;  PRIOR FILING DATE: 2001-03-30
;  PRIOR APPLICATION NUMBER: US 10/276,817
;  PRIOR FILING DATE: 2001-05-16
;  PRIOR APPLICATION NUMBER: US 10/461,673
;  PRIOR FILING DATE: 2003-06-13
;  PRIOR APPLICATION NUMBER: US 10/363,616
;  PRIOR FILING DATE: 2001-08-31
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 1150
;  SOFTWARE: pt_FL_genes Version 6.0
; SEQ ID NO 861
;   LENGTH: 387
;   TYPE: PRT
;   ORGANISM: Homo sapiens
PCT-US04-09202-861

  Query Match           99.1%;  Score 1982;  DB 1;  Length 387;
  Best Local Similarity  98.9%;  Pred. No. 6.9e-171;
  Matches  370;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 MRPGTALQAVLLAVLLVGLRAATGRLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAK 60
              | |||||:|||||||||||: ||||||||||||||||||||||||||||||||||||||
Db         14 MMPGTALKAVLLAVLLVGLQTATGRLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAK 73

Qy         61 EACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         74 EACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWT 133

Qy        121 DGSISQFRNWYVDEPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSD 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        134 DGSISQFRNWYVDEPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSD 193

Qy        181 EKPAVPSREAEGEETELTTPVLPEETQEEDAKKTFKESREAALNLAYILIPSIPLLLLLV 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        194 EKPAVPSREAEGEETELTTPVLPEETQEEDAKKTFKESREAALNLAYILIPSIPLLLLLV 253

Qy        241 VTTVVCWVWICRKRKREQPDPSTKKQHTIWPSPHQGNSPDLEVYNVIRKQSEADLAETRP 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        254 VTTVVCWVWICRKRKREQPDPSTKKQHTIWPSPHQGNSPDLEVYNVIRKQSEADLAETRP 313

Qy        301 DLKNISFRVCSGEATPDDMSCDYDNMAVNPSESGFVTLVSVESGFVTNDIYEFSPDQMGR 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        314 DLKNISFRVCSGEATPDDMSCDYDNMAVNPSESGFVTLVSVESGFVTNDIYEFSPDQMGR 373

Qy        361 SKESGWVENEIYGY 374
              ||||||||||||||
Db        374 SKESGWVENEIYGY 387


RESULT 5
```

```
PCT-US04-09202-1040
; Sequence 1040, Application PC/TUS0409202
; GENERAL INFORMATION:
;  APPLICANT: Zhou, Ping
;  APPLICANT:  Tang, Y. Tom
;  APPLICANT:  Hu, Tianhua
;  APPLICANT:  Wang, Jian-Rui
;  APPLICANT:  Wang, Zhi Wei
;  APPLICANT:  Drmanac, Radoje T.
;  TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
;  FILE REFERENCE: 824CIP/PCT
;  CURRENT APPLICATION NUMBER: PCT/US04/09202
;  CURRENT FILING DATE:  2004-04-06
;  PRIOR APPLICATION NUMBER: US 60/458,824
;  PRIOR FILING DATE: 2003-03-28
;  PRIOR APPLICATION NUMBER: US 10/296,115
;  PRIOR FILING DATE: 2000-12-22
;  PRIOR APPLICATION NUMBER: US 10/275,027
;  PRIOR FILING DATE: 2001-01-25
;  PRIOR APPLICATION NUMBER: US 10/276,774
;  PRIOR FILING DATE: 2001-02-05
;  PRIOR APPLICATION NUMBER: US 10/220,366
;  PRIOR FILING DATE: 2001-02-26
;  PRIOR APPLICATION NUMBER: US 10/221,279
;  PRIOR FILING DATE: 2001-03-05
;  PRIOR APPLICATION NUMBER: US 10/450,763
;  PRIOR FILING DATE: 2001-03-30
;  PRIOR APPLICATION NUMBER: US 10/276,817
;  PRIOR FILING DATE: 2001-05-16
;  PRIOR APPLICATION NUMBER: US 10/461,673
;  PRIOR FILING DATE: 2003-06-13
;  PRIOR APPLICATION NUMBER: US 10/363,616
;  PRIOR FILING DATE: 2001-08-31
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 1150
;  SOFTWARE: pt_FL_genes Version 6.0
; SEQ ID NO 1040
;   LENGTH: 374
;   TYPE: PRT
;   ORGANISM: Homo sapiens
PCT-US04-09202-1040

  Query Match             99.0%;  Score 1979;  DB 1;  Length 374;
  Best Local Similarity   98.7%;  Pred. No. 1.2e-170;
  Matches  369;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 MRPGTALQAVLLAVLLVGLRAATGRLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAK 60
              | ||||||: |||||||||: ||||||||||||||||||||||||||||||||||||||||
Db          1 MMPGTALEGVLLAVLLVGLQTATGRLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAK 60

Qy         61 EACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 EACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWT 120

Qy        121 DGSISQFRNWYVDEPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSD 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 DGSISQFRNWYVDEPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSD 180

Qy        181 EKPAVPSREAEGEETELTTPVLPEETQEEDAKKTFKESREAALNLAYILIPSIPLLLLLV 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 EKPAVPSREAEGEETELTTPVLPEETQEEDAKKTFKESREAALNLAYILIPSIPLLLLLV 240

Qy        241 VTTVVCWVWICRKRKREQPDPSTKKQHTIWPSPHQGNSPDLEVYNVIRKQSEADLAETRP 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 VTTVVCWVWICRKRKREQPDPSTKKQHTIWPSPHQGNSPDLEVYNVIRKQSEADLAETRP 300

Qy        301 DLKNISFRVCSGEATPDDMSCDYDNMAVNPSESGFVTLVSVESGFVTNDIYEFSPDQMGR 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 DLKNISFRVCSGEATPDDMSCDYDNMAVNPSESGFVTLVSVESGFVTNDIYEFSPDQMGR 360

Qy        361 SKESGWVENEIYGY 374
              ||||||||||||||
Db        361 SKESGWVENEIYGY 374
```

```
RESULT 6
PCT-US04-09202-363
; Sequence 363, Application PC/TUS0409202
; GENERAL INFORMATION:
;  APPLICANT: Zhou, Ping
;  APPLICANT:  Tang, Y. Tom
;  APPLICANT:  Hu, Tianhua
;  APPLICANT:  Wang, Jian-Rui
;  APPLICANT:  Wang, Zhi Wei
;  APPLICANT:  Drmanac, Radoje T.
;  TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
;  FILE REFERENCE: 824CIP/PCT
;  CURRENT APPLICATION NUMBER: PCT/US04/09202
;  CURRENT FILING DATE:  2004-04-06
;  PRIOR APPLICATION NUMBER: US 60/458,824
;  PRIOR FILING DATE: 2003-03-28
;  PRIOR APPLICATION NUMBER: US 10/296,115
;  PRIOR FILING DATE: 2000-12-22
;  PRIOR APPLICATION NUMBER: US 10/275,027
;  PRIOR FILING DATE: 2001-01-25
;  PRIOR APPLICATION NUMBER: US 10/276,774
;  PRIOR FILING DATE: 2001-02-05
;  PRIOR APPLICATION NUMBER: US 10/220,366
;  PRIOR FILING DATE: 2001-02-26
;  PRIOR APPLICATION NUMBER: US 10/221,279
;  PRIOR FILING DATE: 2001-03-05
;  PRIOR APPLICATION NUMBER: US 10/450,763
;  PRIOR FILING DATE: 2001-03-30
;  PRIOR APPLICATION NUMBER: US 10/276,817
;  PRIOR FILING DATE: 2001-05-16
;  PRIOR APPLICATION NUMBER: US 10/461,673
;  PRIOR FILING DATE: 2003-06-13
;  PRIOR APPLICATION NUMBER: US 10/363,616
;  PRIOR FILING DATE: 2001-08-31
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 1150
;  SOFTWARE: pt_FL_genes Version 6.0
; SEQ ID NO 363
;   LENGTH: 356
;   TYPE: PRT
;   ORGANISM: Homo sapiens
PCT-US04-09202-363

  Query Match             90.2%;  Score 1805;  DB 1;  Length 356;
  Best Local Similarity   97.1%;  Pred. No. 6.4e-155;
  Matches  334;  Conservative    2;  Mismatches    8;  Indels    0;  Gaps    0;

Qy         31 PVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLL 90
              || |||||||| |||||||||||||| : |||||||||:   ||||||||||||||||
Db         13 PVRPGGTQRPCYKAIYFHDTSRRLNFEEAHKKCRRDGGQLVTARPEDEQKLIEKFIENLL 72

Qy         91 PSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEPSCGSEVCVVMYHQPS 150
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         73 PSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEPSCGSEVCVVMYHQPS 132

Qy        151 APAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGEETELTTPVLPEETQEED 210
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        133 APAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGEETELTTPVLPEETQEED 192

Qy        211 AKKTFKESREAALNLAYILIPSIPLLLLLVVTTVVCWVWICRKRKREQPDPSTKKQHTIW 270
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        193 AKKTFKESREAALNLAYILIPSIPLLLLLVVTTVVCWVWICRKRKREQPDPSTKKQHTIW 252

Qy        271 PSPHQGNSPDLEVYNVIRKQSEADLAETRPDLKNISFRVCSGEATPDDMSCDYDNMAVNP 330
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        253 PSPHQGNSPDLEVYNVIRKQSEADLAETRPDLKNISFRVCSGEATPDDMSCDYDNMAVNP 312

Qy        331 SESGFVTLVSVESGFVTNDIYEFSPDQMGRSKESGWVENEIYGY 374
              ||||||||||||||||||||||||||||||||||||||||||||
Db        313 SESGFVTLVSVESGFVTNDIYEFSPDQMGRSKESGWVENEIYGY 356
```

```
RESULT 7
US-10-170-205E-35136
; Sequence 35136, Application US/10170205E
; GENERAL INFORMATION:
;  APPLICANT: ADAMS, Mark
;  TITLE OF INVENTION: DEVICES, SUCH AS ARRAYS, COMPRISED OF HUMAN PROTEINS OR PROTEIN
;  TITLE OF INVENTION:  CAPTURE AGENTS, AND USES THEREOF
;  FILE REFERENCE: CL001381
;  CURRENT APPLICATION NUMBER: US/10/170,205E
;  CURRENT FILING DATE:  2002-06-13
;  NUMBER OF SEQ ID NOS: 40312
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 35136
;   LENGTH: 358
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-170-205E-35136

  Query Match             89.2%;  Score 1784;  DB 6;  Length 358;
  Best Local Similarity   94.1%;  Pred. No. 5.1e-153;
  Matches  337;  Conservative    1;  Mismatches    8;  Indels   12;  Gaps    1;

Qy         29 GQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIEN 88
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIEN 60

Qy         89 LLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEPSCGSEVCVVMYHQ 148
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 LLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEPSCGSEVCVVMYHQ 120

Qy        149 PSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGEETELTTPVLPEETQE 208
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 PSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGEETELTTPVLPEETQE 180

Qy        209 EDAKKTFKESREAALNLAYILIPSIPLLLLLVVTTVVCWVW----------ICRKRKR 256
              |||||||||||||||||||||||||||||||||||||    :                |||
Db        181 EDAKKTFKESREAALNLAYILIPSIPLLLLLVVTTVEAQGYSQSPGLKHSALVFFLPRKR 240

Qy        257 EQPDPSTKKQHTIWPSPHQGNSPDLEVYNVIRKQSEADLAETRPDLKNISFRVCSGEATP 316
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 EQPDPSTKKQHTIWPSPHQGNSPDLEVYNVIRKQSEADLAETRPDLKNISFRVCSGEATP 300

Qy        317 DDMSCDYDNMAVNPSESGFVTLVSVESGFVTNDIYEFSPDQMGRSKESGWVENEIYGY 374
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 DDMSCDYDNMAVNPSESGFVTLVSVESGFVTNDIYEFSPDQMGRSKESGWVENEIYGY 358


RESULT 8
PCT-US04-09202-1010
; Sequence 1010, Application PC/TUS0409202
; GENERAL INFORMATION:
;  APPLICANT: Zhou, Ping
;  APPLICANT:  Tang, Y. Tom
;  APPLICANT:  Hu, Tianhua
;  APPLICANT:  Wang, Jian-Rui
;  APPLICANT:  Wang, Zhi Wei
;  APPLICANT:  Drmanac, Radoje T.
;  TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
;  FILE REFERENCE: 824CIP/PCT
;  CURRENT APPLICATION NUMBER: PCT/US04/09202
;  CURRENT FILING DATE:  2004-04-06
;  PRIOR APPLICATION NUMBER: US 60/458,824
;  PRIOR FILING DATE: 2003-03-28
;  PRIOR APPLICATION NUMBER: US 10/296,115
;  PRIOR FILING DATE: 2000-12-22
;  PRIOR APPLICATION NUMBER: US 10/275,027
;  PRIOR FILING DATE: 2001-01-25
;  PRIOR APPLICATION NUMBER: US 10/276,774
;  PRIOR FILING DATE: 2001-02-05
;  PRIOR APPLICATION NUMBER: US 10/220,366
;  PRIOR FILING DATE: 2001-02-26
;  PRIOR APPLICATION NUMBER: US 10/221,279
;  PRIOR FILING DATE: 2001-03-05
;  PRIOR APPLICATION NUMBER: US 10/450,763
;  PRIOR FILING DATE: 2001-03-30
;  PRIOR APPLICATION NUMBER: US 10/276,817
;  PRIOR FILING DATE: 2001-05-16
;  PRIOR APPLICATION NUMBER: US 10/461,673
;  PRIOR FILING DATE: 2003-06-13
;  PRIOR APPLICATION NUMBER: US 10/363,616
;  PRIOR FILING DATE: 2001-08-31
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 1150
;  SOFTWARE: pt_FL_genes Version 6.0
; SEQ ID NO 1010
;   LENGTH: 389
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(389)
;   OTHER INFORMATION: Xaa = any amino acid or nothing
PCT-US04-09202-1010

  Query Match             88.3%;  Score 1765.5;  DB 1;  Length 389;
  Best Local Similarity   93.4%;  Pred. No. 2.7e-151;
  Matches  327;  Conservative    9;  Mismatches   13;  Indels    1;  Gaps    1;

Qy         26 LLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEKF 85
              |  |||||||||||||||||||||||||||||||| | :||||||||  ||||||||||||||
Db         40 LRGGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEVKFSCRRDGGQLGSIESEDEQKLIEKF 99

Qy         86 IENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEPSCGSEVCVVM 145
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        100 IENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEPSCGSEVCVVM 159

Qy        146 YHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGEETELTTPVLPEE 205
              |||||||||||||||||||||||||||||||||||||||||||||||| : |||||||||||||
Db        160 YHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSRRSXGEETELTTPVLPEE 219

Qy        206 TQEEDAKKTFKESREAALNLAYILIPSIPLLLLLVVTTVVCWVWICRKRKREQPDPSTKK 265
              |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db        220 TQEEDAKKTFKESREAALNLAYILIPSIPLLLLLVVTTSVCWVWICRKRKREQPDPSTKK 279

Qy        266 QHTIWPSPHQGNSPDLEVYNVIRK-QSEADLAETRPDLKNISFRVCSGEATPDDMSCDYD 324
              |||||||||||||||||||| :|  ::: |:||||||:|||||||| |: |||||||||
Db        280 QHTIWPSPHQGNSPDLEVYNVSKKTNAKSFLSETRPDLRNISFRVCSEESPPDDMSCDYD 339

Qy        325 NMAVNPSESGFVTLVSVESGFVTNDIYEFSPDQMGRSKESGWVENEIYGY 374
              ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        340 NMAVNPSESGFVTLVSVESGFVTNDIYEFSPDQMGRSKESGWVENEIYGY 389


RESULT 9
US-10-897-911-3
; Sequence 3, Application US/10897911
; GENERAL INFORMATION:
;  APPLICANT: Daniel E.H. Afar
;  APPLICANT:  Rene S. Hubert
;  APPLICANT:  Aya Jakobovits
;  APPLICANT:  Arthur B. Raitano
;  TITLE OF INVENTION: C-TYPE LECTIN TRANSMEMBRANE ANTIGEN EXPRESSED IN HUMAN PROSTATE C.
;  TITLE OF INVENTION:  USES THEREOF
;  FILE REFERENCE: 51158-20020.12
;  CURRENT APPLICATION NUMBER: US/10/897,911
;  CURRENT FILING DATE:  2004-07-23
;  PRIOR APPLICATION NUMBER: 10/460,512
;  PRIOR FILING DATE: 2003-06-11
;  PRIOR APPLICATION NUMBER: 09/638,203
;  PRIOR FILING DATE: 2000-08-11
;  PRIOR APPLICATION NUMBER: 60/148,935
;  PRIOR FILING DATE: 1999-08-12
;  NUMBER OF SEQ ID NOS: 47
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 3
```

```
;   LENGTH: 260
;   TYPE: PRT
;   ORGANISM: Hamster
US-10-897-911-3

  Query Match             59.9%;  Score 1198.5;  DB 6;  Length 260;
  Best Local Similarity   84.2%;  Pred. No. 3.3e-100;
  Matches  219;  Conservative   17;  Mismatches   23;  Indels    1;  Gaps    1;

Qy         25 RLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEK 84
              |||||| |||||||:|||||||||||  :|||||||||||||||||||||:||||:||||
Db          1 RLLSGQLVCRGGTRRPCYKVIYFHDAFQRLNFEEAKEACRRDGGQLVSIETEDEQRLIEK 60

Qy         85 FIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEPSCGSEVCVV 144
              |||||| |||||||||||  | || |:|||||||||||||| |||||||||||||||||||
Db         61 FIENLLASDGDFWIGLRRLEVKQVNNTACQDLYAWTDGSTSQFRNWYVDEPSCGSEVCVV 120

Qy        145 MYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPA-VPSREAEGEETELTTPVLP 203
              |||||||| |||| |||||||||||||||||||:||||:  ||     || ||  |||||
Db        121 MYHQPSAPPGIGGSYMFQWNDDRCNMKNNFICKYADEKPSTTPSIRPGGEATEPPTPVLP 180

Qy        204 EETQEEDAKKTFKESREAALNLAYILIPSIPLLLLLVVTTVVCWVWICRKRKREQPDPST 263
              ||||:|| |:|||||||||||||||||||||| ||||||:  |||||||:||:|||||:|
Db        181 EETQKEDTKETFKESREAALNLAYILIPSIPLFLLLVVTSAACWVWICRRRKQEQPDPTT 240

Qy        264 KKQHTIWPSPHQGNSPDLEV 283
              |:||||||:||| |||:|:|
Db        241 KEQHTIWPTPHQENSPNLDV 260


RESULT 10
US-10-898-615-3
; Sequence 3, Application US/10898615
; GENERAL INFORMATION:
;  APPLICANT: Daniel E.H. Afar
;  APPLICANT:  Rene S. Hubert
;  APPLICANT:  Aya Jakobovits
;  APPLICANT:  Arthur B. Raitano
;  TITLE OF INVENTION: C-TYPE LECTIN TRANSMEMBRANE ANTIGEN EXPRESSED IN HUMAN PROSTATE C
;  TITLE OF INVENTION:  USES THEREOF
;  FILE REFERENCE: 51158-20020.11
;  CURRENT APPLICATION NUMBER: US/10/898,615
;  CURRENT FILING DATE:  2004-07-23
;  PRIOR APPLICATION NUMBER: 10/460,512
;  PRIOR FILING DATE: 2003-06-11
;  PRIOR APPLICATION NUMBER: 09/638,203
;  PRIOR FILING DATE: 2000-08-11
;  PRIOR APPLICATION NUMBER: 60/148,935
;  PRIOR FILING DATE: 1999-08-12
;  NUMBER OF SEQ ID NOS: 47
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 3
;   LENGTH: 260
;   TYPE: PRT
;   ORGANISM: Hamster
US-10-898-615-3

  Query Match             59.9%;  Score 1198.5;  DB 6;  Length 260;
  Best Local Similarity   84.2%;  Pred. No. 3.3e-100;
  Matches  219;  Conservative   17;  Mismatches   23;  Indels    1;  Gaps    1;

Qy         25 RLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEK 84
              |||||| |||||||:|||||||||||  :|||||||||||||||||||||:||||:||||
Db          1 RLLSGQLVCRGGTRRPCYKVIYFHDAFQRLNFEEAKEACRRDGGQLVSIETEDEQRLIEK 60

Qy         85 FIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEPSCGSEVCVV 144
              |||||| |||||||||||  | || |:|||||||||||||| |||||||||||||||||||
Db         61 FIENLLASDGDFWIGLRRLEVKQVNNTACQDLYAWTDGSTSQFRNWYVDEPSCGSEVCVV 120

Qy        145 MYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPA-VPSREAEGEETELTTPVLP 203
              |||||||| |||| |||||||||||||||||||:||||:  ||     || ||  |||||
Db        121 MYHQPSAPPGIGGSYMFQWNDDRCNMKNNFICKYADEKPSTTPSIRPGGEATEPPTPVLP 180

Qy        204 EETQEEDAKKTFKESREAALNLAYILIPSIPLLLLLVVTTVVCWVWICRKRKREQPDPST 263
              ||||:|| |:|||||||||||||||||||||| ||||||:  |||||||:||:|||||:|
Db        181 EETQKEDTKETFKESREAALNLAYILIPSIPLFLLLVVTSAACWVWICRRRKQEQPDPTT 240

Qy        264 KKQHTIWPSPHQGNSPDLEV 283
              |:||||||:||| |||:|:|
Db        241 KEQHTIWPTPHQENSPNLDV 260


RESULT 11
US-10-152-372-540
; Sequence 540, Application US/10152372
; GENERAL INFORMATION:
;  APPLICANT: Baker,Kevin P.
;  APPLICANT:  Beresini,Maureen
;  APPLICANT:  DeForge,Laura
;  APPLICANT:  Desnoyers,Luc
;  APPLICANT:  Filvaroff,Ellen
;  APPLICANT:  Gao,Wei-Qiang
;  APPLICANT:  Gerritsen,Mary E.
;  APPLICANT:  Goddard,Audrey
;  APPLICANT:  Godowski,Paul J.
;  APPLICANT:  Gurney,Austin L.
;  APPLICANT:  Sherwood,Steven
;  APPLICANT:  Smith,Victoria
;  APPLICANT:  Stewart,Timothy A.
;  APPLICANT:  Tumas,Daniel
;  APPLICANT:  Watanabe,Colin K
;  APPLICANT:  Wood,William
;  APPLICANT:  Zhang, Zemin
;  TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
;  TITLE OF INVENTION:  ACIDS ENCODING THE SAME
;  FILE REFERENCE: P3330R1C394
;  CURRENT APPLICATION NUMBER: US/10/152,372
;  CURRENT FILING DATE:  2002-05-21
;  PRIOR APPLICATION NUMBER: 60/049911
;  PRIOR FILING DATE: 1997-06-18
;  PRIOR APPLICATION NUMBER: 60/056974
;  PRIOR FILING DATE: 1997-08-26
;  PRIOR APPLICATION NUMBER: 60/059113
;  PRIOR FILING DATE: 1997-09-17
;  PRIOR APPLICATION NUMBER: 60/059115
;  PRIOR FILING DATE: 1997-09-17
;  PRIOR APPLICATION NUMBER: 60/059117
;  PRIOR FILING DATE: 1997-09-17
;  PRIOR APPLICATION NUMBER: 60/059122
;  PRIOR FILING DATE: 1997-09-17
;  PRIOR APPLICATION NUMBER: 60/059184
;  PRIOR FILING DATE: 1997-09-17
;  PRIOR APPLICATION NUMBER: 60/059263
;  PRIOR FILING DATE: 1997-09-18
;  PRIOR APPLICATION NUMBER: 60/059352
;  PRIOR FILING DATE: 1997-09-19
;  PRIOR APPLICATION NUMBER: 60/059588
;  PRIOR FILING DATE: 1997-09-19
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 540
;   LENGTH: 273
;   TYPE: PRT
;   ORGANISM: Homo Sapien
US-10-152-372-540

  Query Match             32.2%;  Score 645;  DB 6;  Length 273;
  Best Local Similarity   48.4%;  Pred. No. 4.4e-50;
  Matches  137;  Conservative   40;  Mismatches   80;  Indels   26;  Gaps    9;

Qy         10 VLLAVLLVGLRAATGRLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQ 69
              :| | || |  |   |::||| ||    : ||||: |||: | |::|:||: ||  :||
Db          8 LLGAALLCGHGAFCRRVVSGQKVCFADFKHPCYKMAYFHELSSRVSFQEARLACESEGGV 67

Qy         70 LVSIESEDEQKLIEKFIENLLP-----SDGDFWIGLRRREEKQSNSTACQDLYAWTDGSI 124
              |:|:|:| ||||||  ::||       |||||||||  |  : |: | || ||| |:|||
```

```
Db         68 LLSLENEAEQKLIESMLQNLTKPGTGISDGDFWIGLWRNGDGQT-SGACPDLYQWSDGSN 126

Qy        125 SQFRNWYVDEPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDE-KP 183
              ||:|||| |||||||| ||||||||:|  |:||||::||||||||||:|:||||  |  |
Db        127 SQYRNWYTDEPSCGSEKCVVMYHQPTANPGLGGPYLYQWNDDRCNMKHNYICKYEPEINP 186

Qy        184 AVPSREAEGEETELTTPVLPEETQEEDAKKTFKESREAAL--NLAYILIPSIPLLLLLVV 241
                |      |:  ||    | :| :           || :  || |::||:||||||::|
Db        187 TAPV-----EKPYLTNQ--PGDTHQNVV------VTEAGIIPNLIYVVIPTIPLLLLILV 233

Qy        242 TTVVCWVWICRKRK-REQPDPSTKKQHTIWPSPHQGNSPDLEV 283
                  |   :  | | | :  |:   | |:| |        :||
Db        234 AFGTCCFQMLHKSKGRTKTSPN---QSTLWISKSTRKESGMEV 273
```

RESULT 12
US-10-123-155-540
; Sequence 540, Application US/10123155
; GENERAL INFORMATION:
; APPLICANT: Baker,Kevin P.
; APPLICANT: Beresini,Maureen
; APPLICANT: DeForge,Laura
; APPLICANT: Desnoyers,Luc
; APPLICANT: Filvaroff,Ellen
; APPLICANT: Gao,Wei-Qiang
; APPLICANT: Gerritsen,Mary E.
; APPLICANT: Goddard,Audrey
; APPLICANT: Godowski,Paul J.
; APPLICANT: Gurney,Austin L.
; APPLICANT: Sherwood,Steven
; APPLICANT: Smith,Victoria
; APPLICANT: Stewart,Timothy A.
; APPLICANT: Tumas,Daniel
; APPLICANT: Watanabe,Colin K
; APPLICANT: Wood,William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3330R1C30
; CURRENT APPLICATION NUMBER: US/10/123,155
; CURRENT FILING DATE: 2002-04-15
; PRIOR APPLICATION NUMBER: 60/049911
; PRIOR FILING DATE: 1997-06-18
; PRIOR APPLICATION NUMBER: 60/056974
; PRIOR FILING DATE: 1997-08-26
; PRIOR APPLICATION NUMBER: 60/059113
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059115
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059117
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059122
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059184
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059352
; PRIOR FILING DATE: 1997-09-19
; PRIOR APPLICATION NUMBER: 60/059588
; PRIOR FILING DATE: 1997-09-19
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 540
; LENGTH: 273
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-123-155-540

Query Match 32.2%; Score 645; DB 6; Length 273;
Best Local Similarity 48.4%; Pred. No. 4.4e-50;
Matches 137; Conservative 40; Mismatches 80; Indels 26; Gaps 9;

```
Qy         10 VLLAVLLVGLRAATGRLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQ 69
              :| | || |  |    |::||| ||    : ||||: |||: | |::|:||: ||  :||
Db          8 LLGAALLCGHGAFCRRVVSGQKVCFADFKHPCYKMAYFHELSSRVSFQEARLACESEGGV 67

Qy         70 LVSIESEDEQKLIEKFIENLLP-----SDGDFWIGLRRREEKQSNSTACQDLYAWTDGSI 124
              |:|:|:| ||||||  ::||        |||||||||  |  : |: | || ||| |:|||
Db         68 LLSLENEAEQKLIESMLQNLTKPGTGISDGDFWIGLWRNGDGQT-SGACPDLYQWSDGSN 126

Qy        125 SQFRNWYVDEPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDE-KP 183
              ||:|||| |||||||| ||||||||:|  |:||||::||||||||||:|:||||  |  |
Db        127 SQYRNWYTDEPSCGSEKCVVMYHQPTANPGLGGPYLYQWNDDRCNMKHNYICKYEPEINP 186

Qy        184 AVPSREAEGEETELTTPVLPEETQEEDAKKTFKESREAAL--NLAYILIPSIPLLLLLVV 241
                |      |:  ||    | :| :           || :  || |::||:||||||::|
Db        187 TAPV-----EKPYLTNQ--PGDTHQNVV------VTEAGIIPNLIYVVIPTIPLLLLILV 233

Qy        242 TTVVCWVWICRKRK-REQPDPSTKKQHTIWPSPHQGNSPDLEV 283
                  |   :  | | | :  |:   | |:| |        :||
Db        234 AFGTCCFQMLHKSKGRTKTSPN---QSTLWISKSTRKESGMEV 273
```

RESULT 13
US-10-063-685-168
; Sequence 168, Application US/10063685
; GENERAL INFORMATION:
; APPLICANT: Eaton,Dan L.
; APPLICANT: Filvaroff,Ellen
; APPLICANT: Gerritsen,Mary E.
; APPLICANT: Goddard,Audrey
; APPLICANT: Godowski,Paul J.
; APPLICANT: Grimaldi,Christopher J.
; APPLICANT: Gurney,Austin L.
; APPLICANT: Watanabe,Colin K.
; APPLICANT: Wood,William I.
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3230R1C1
; CURRENT APPLICATION NUMBER: US/10/063,685
; CURRENT FILING DATE: 2002-05-08
; PRIOR APPLICATION NUMBER: 60/063435
; PRIOR FILING DATE: 1997-10-29
; PRIOR APPLICATION NUMBER: 60/064215
; PRIOR FILING DATE: 1997-10-29
; PRIOR APPLICATION NUMBER: 60/082797
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/083495
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/085579
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/087759
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/088021
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088029
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088030
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088734
; PRIOR FILING DATE: 1998-06-10
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 170
; SEQ ID NO 168
; LENGTH: 273
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-063-685-168

Query Match 32.2%; Score 645; DB 6; Length 273;
Best Local Similarity 48.4%; Pred. No. 4.4e-50;
Matches 137; Conservative 40; Mismatches 80; Indels 26; Gaps 9;

```
Qy         10 VLLAVLLVGLRAATGRLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQ 69
              :| | || |  |    |::||| ||    : ||||: |||: | |::|:||: ||  :||
```

```
Db          8 LLGAALLCGHGAFCRRVVSGQKVCFADFKHPCYKMAYFHELSSRVSFQEARLACESEGGV 67

Qy         70 LVSIESEDEQKLIEKFIENLLP-----SDGDFWIGLRRREEKQSNSTACQDLYAWTDGSI 124
              |:|:|:| ||||||  ::||       |||||||||| |  : |: | || ||| |:|||
Db         68 LLSLENEAEQKLIESMLQNLTKPGTGISDGDFWIGLWRNGDGQT-SGACPDLYQWSDGSN 126

Qy        125 SQFRNWYVDEPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDE-KP 183
              ||:|||| |||||||| ||||||||:|  |:||||::||||||||||:|:||||  |  |
Db        127 SQYRNWYTDEPSCGSEKCVVMYHQPTANPGLGGPYLYQWNDDRCNMKHNYICKYEPEINP 186

Qy        184 AVPSREAEGEETELTTPVLPEETQEEDAKKTFKESREAAL--NLAYILIPSIPLLLLLVV 241
                |      |:  ||    | :| :           || :  || |::||:||||||::|
Db        187 TAPV-----EKPYLTNQ--PGDTHQNVV------VTEAGIIPNLIYVVIPTIPLLLLILV 233

Qy        242 TTVVCWVWICRKRK-REQPDPSTKKQHTIWPSPHQGNSPDLEV 283
                 |   :  | | | :  |:   | |:| |        :||
Db        234 AFGTCCFQMLHKSKGRTKTSPN---QSTLWISKSTRKESGMEV 273


RESULT 14
US-10-143-117-540
; Sequence 540, Application US/10143117
; GENERAL INFORMATION:
;  APPLICANT: Baker,Kevin P.
;  APPLICANT:  Beresini,Maureen
;  APPLICANT:  DeForge,Laura
;  APPLICANT:  Desnoyers,Luc
;  APPLICANT:  Filvaroff,Ellen
;  APPLICANT:  Gao,Wei-Qiang
;  APPLICANT:  Gerritsen,Mary E.
;  APPLICANT:  Goddard,Audrey
;  APPLICANT:  Godowski,Paul J.
;  APPLICANT:  Gurney,Austin L.
;  APPLICANT:  Sherwood,Steven
;  APPLICANT:  Smith,Victoria
;  APPLICANT:  Stewart,Timothy A.
;  APPLICANT:  Tumas,Daniel
;  APPLICANT:  Watanabe,Colin K
;  APPLICANT:  Wood,William
;  APPLICANT:  Zhang, Zemin
;  TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
;  TITLE OF INVENTION:  ACIDS ENCODING THE SAME
;  FILE REFERENCE: P3330R1C221
;  CURRENT APPLICATION NUMBER: US/10/143,117
;  CURRENT FILING DATE:  2002-05-09
;  PRIOR APPLICATION NUMBER: 60/049911
;  PRIOR FILING DATE: 1997-06-18
;  PRIOR APPLICATION NUMBER: 60/056974
;  PRIOR FILING DATE: 1997-08-26
;  PRIOR APPLICATION NUMBER: 60/059113
;  PRIOR FILING DATE: 1997-09-17
;  PRIOR APPLICATION NUMBER: 60/059115
;  PRIOR FILING DATE: 1997-09-17
;  PRIOR APPLICATION NUMBER: 60/059117
;  PRIOR FILING DATE: 1997-09-17
;  PRIOR APPLICATION NUMBER: 60/059122
;  PRIOR FILING DATE: 1997-09-17
;  PRIOR APPLICATION NUMBER: 60/059184
;  PRIOR FILING DATE: 1997-09-17
;  PRIOR APPLICATION NUMBER: 60/059263
;  PRIOR FILING DATE: 1997-09-18
;  PRIOR APPLICATION NUMBER: 60/059352
;  PRIOR FILING DATE: 1997-09-19
;  PRIOR APPLICATION NUMBER: 60/059588
;  PRIOR FILING DATE: 1997-09-19
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 540
;   LENGTH: 273
;   TYPE: PRT
;   ORGANISM: Homo Sapien
US-10-143-117-540

  Query Match             32.2%;  Score 645;  DB 6;  Length 273;
  Best Local Similarity   48.4%;  Pred. No. 4.4e-50;
  Matches  137;  Conservative   40;  Mismatches   80;  Indels   26;  Gaps    9;

Qy         10 VLLAVLLVGLRAATGRLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQ 69
              :| | || |  |   |::||| ||    : ||||: |||: | |::|:||: ||  :||
Db          8 LLGAALLCGHGAFCRRVVSGQKVCFADFKHPCYKMAYFHELSSRVSFQEARLACESEGGV 67

Qy         70 LVSIESEDEQKLIEKFIENLLP-----SDGDFWIGLRRREEKQSNSTACQDLYAWTDGSI 124
              |:|:|:| ||||||  ::||       |||||||||| |  : |: | || ||| |:|||
Db         68 LLSLENEAEQKLIESMLQNLTKPGTGISDGDFWIGLWRNGDGQT-SGACPDLYQWSDGSN 126

Qy        125 SQFRNWYVDEPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDE-KP 183
              ||:|||| |||||||| ||||||||:|  |:||||::||||||||||:|:||||  |  |
Db        127 SQYRNWYTDEPSCGSEKCVVMYHQPTANPGLGGPYLYQWNDDRCNMKHNYICKYEPEINP 186

Qy        184 AVPSREAEGEETELTTPVLPEETQEEDAKKTFKESREAAL--NLAYILIPSIPLLLLLVV 241
                |      |:  ||    | :| :           || :  || |::||:||||||::|
Db        187 TAPV-----EKPYLTNQ--PGDTHQNVV------VTEAGIIPNLIYVVIPTIPLLLLILV 233

Qy        242 TTVVCWVWICRKRK-REQPDPSTKKQHTIWPSPHQGNSPDLEV 283
                 |   :  | | | :  |:   | |:| |        :||
Db        234 AFGTCCFQMLHKSKGRTKTSPN---QSTLWISKSTRKESGMEV 273


RESULT 15
US-10-127-823A-540
; Sequence 540, Application US/10127823A
; GENERAL INFORMATION:
;  APPLICANT: Baker,Kevin P.
;  APPLICANT:  Beresini,Maureen
;  APPLICANT:  DeForge,Laura
;  APPLICANT:  Desnoyers,Luc
;  APPLICANT:  Filvaroff,Ellen
;  APPLICANT:  Gao,Wei-Qiang
;  APPLICANT:  Gerritsen,Mary E.
;  APPLICANT:  Goddard,Audrey
;  APPLICANT:  Godowski,Paul J.
;  APPLICANT:  Gurney,Austin L.
;  APPLICANT:  Sherwood,Steven
;  APPLICANT:  Smith,Victoria
;  APPLICANT:  Stewart,Timothy A.
;  APPLICANT:  Tumas,Daniel
;  APPLICANT:  Watanabe,Colin K
;  APPLICANT:  Wood,William
;  APPLICANT:  Zhang, Zemin
;  TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
;  TITLE OF INVENTION:  ACIDS ENCODING THE SAME
;  FILE REFERENCE: 39780-3330R1C189
;  CURRENT APPLICATION NUMBER: US/10/127,823A
;  CURRENT FILING DATE:  2002-04-22
;  PRIOR APPLICATION NUMBER: US 10/028,072
;  PRIOR FILING DATE: 2001-12-19
;  PRIOR APPLICATION NUMBER: PCT/US00/32678
;  PRIOR FILING DATE: 2000-12-01
;  PRIOR APPLICATION NUMBER: US 60/170,262
;  PRIOR FILING DATE: 1999-12-09
;  NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 540
;   LENGTH: 273
;   TYPE: PRT
;   ORGANISM: Homo Sapien
US-10-127-823A-540

  Query Match             32.2%;  Score 645;  DB 6;  Length 273;
  Best Local Similarity   48.4%;  Pred. No. 4.4e-50;
  Matches  137;  Conservative   40;  Mismatches   80;  Indels   26;  Gaps    9;

Qy         10 VLLAVLLVGLRAATGRLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQ 69
              :| | || |  |   |::||| ||    : ||||: |||: | |::|:||: ||  :||
Db          8 LLGAALLCGHGAFCRRVVSGQKVCFADFKHPCYKMAYFHELSSRVSFQEARLACESEGGV 67

Qy         70 LVSIESEDEQKLIEKFIENLLP-----SDGDFWIGLRRREEKQSNSTACQDLYAWTDGSI 124
```

```
                |:|:|:| ||||||  ::||       |||||||||  |  :  |:  |  ||  |||  |:|||
Db          68 LLSLENEAEQKLIESMLQNLTKPGTGISDGDFWIGLWRNGDGQT-SGACPDLYQWSDGSN 126

Qy         125 SQFRNWYVDEPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDE-KP 183
               ||:|||| ||||||| ||||||||:|  |:||||::|||||||||||:|:||||  |  |
Db         127 SQYRNWYTDEPSCGSEKCVVMYHQPTANPGLGGPYLYQWNDDRCNMKHNYICKYEPEINP 186

Qy         184 AVPSREAEGEETELTTPVLPEETQEEDAKKTFKESREAAL--NLAYILIPSIPLLLLLVV 241
                 |      |:  ||    | :| :           || :  ||  |::||:||||||::|
Db         187 TAPV-----EKPYLTNQ--PGDTHQNVV------VTEAGIIPNLIYVVIPTIPLLLLILV 233

Qy         242 TTVVCWVWICRKRK-REQPDPSTKKQHTIWPSPHQGNSPDLEV 283
                   |   :  | | | :  |:   | |:| |        :||
Db         234 AFGTCCFQMLHKSKGRTKTSPN---QSTLWISKSTRKESGMEV 273
```

Search completed: September 9, 2004, 22:31:52
Job time : 64 secs

This Page Blank (uspto)

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: September 9, 2004, 22:20:01 ; Search time 20 Seconds
(without alignments)
1798.782 Million cell updates/sec

Title: US-09-887-855-2
Perfect score: 2000
Sequence: 1 MRPGTALQAVLLAVLLVGLR.....PDQMGRSKESGWVENEIYGY 374

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_78:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	183	9.2	1456	1	A36563	mannose receptor p
2	178.5	8.9	1455	1	A48925	mannose receptor p
3	177	8.8	1268	2	S52781	neurocan - mouse
4	174.5	8.7	1643	2	T14274	versican precursor
5	174.5	8.7	3381	2	T42389	versican precursor
6	174	8.7	1257	2	S28764	neurocan precursor
7	174	8.7	2397	1	A55535	versican precursor
8	174	8.7	2409	1	A60979	versican precursor
9	171	8.6	3562	2	A47171	chondroitin sulfat
10	170.5	8.5	1479	2	T42710	mannose receptor,
11	158.5	7.9	1340	2	A39808	proteoglycan core
12	158.5	7.9	2327	2	T42630	aggrecan - bovine
13	158.5	7.9	2415	1	A39086	aggrecan precursor
14	156	7.8	162	1	LNRC1	lectin BRA3-1 prec
15	154.5	7.7	612	2	B42755	E-selectin precurs
16	153.5	7.7	2124	2	A28452	proteoglycan core
17	152	7.6	912	2	A54423	brevican precursor
18	151	7.5	162	1	LNRC3	lectin BRA3-2 prec
19	150.5	7.5	459	2	T24425	hypothetical prote
20	150	7.5	330	2	T46256	brevican - human (
21	149	7.4	321	1	LNHUER	IgE Fc receptor II
22	148.5	7.4	152	2	JC4690	coagulation factor
23	148.5	7.4	404	2	A46274	HIV gp120-binding
24	148.5	7.4	2132	1	A55182	aggrecan precursor
25	148	7.4	253	2	E89130	protein F52E1.2 [i
26	147	7.3	883	2	S57653	brevican precursor
27	146	7.3	2109	1	I50421	aggrecan precursor
28	145.5	7.3	742	2	JC7595	scavenger receptor
29	145	7.2	883	2	S49126	brevican precursor
30	144.5	7.2	173	2	S10548	lectin - barnacle
31	144.5	7.2	372	2	S23936	L-selectin precurs
32	143.5	7.2	129	2	JC4329	coagulation factor
33	143.5	7.2	372	1	A32375	L-selectin precurs
34	143.5	7.2	463	2	T26655	hypothetical prote
35	143	7.1	321	2	T26152	hypothetical prote
36	142.5	7.1	131	2	JC5058	bitiscetin alpha c
37	142	7.1	1487	2	S48719	phospholipase-A(2)
38	141.5	7.1	165	2	A47148	reg I, regeneratin
39	141.5	7.1	331	1	LNMSER	IgE Fc receptor, l
40	140	7.0	370	2	S22124	L-selectin precurs
41	139	7.0	248	1	LNHUPS	pulmonary surfacta
42	139	7.0	248	1	LNHUP6	pulmonary surfacta
43	139	7.0	248	1	LNHUP1	pulmonary surfacta
44	139	7.0	283	1	LNFHLS	lectin precursor -
45	138.5	6.9	172	2	S32489	lectin - Iberian r

ALIGNMENTS

RESULT 1
A36563
mannose receptor precursor - human
C;Species: Homo sapiens (man)
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C;Accession: A36563; A60926; A44255; B44255; C44255; D44255; E44255; F44255; G44255; H44
R;Taylor, M.E.; Conary, J.T.; Lennartz, M.R.; Stahl, P.D.; Drickamer, K.
J. Biol. Chem. 265, 12156-12162, 1990
A;Title: Primary structure of the mannose receptor contains multiple motifs resembling c
A;Reference number: A36563; MUID:90324192; PMID:2373685
A;Accession: A36563
A;Molecule type: mRNA
A;Residues: 1-1456 <TAY>
A;Cross-references: GB:J05550; NID:g188675; PIDN:AAA59868.1; PID:g188676
A;Note: parts of this sequence, including the amino end of the mature protein, were conf
R;Ezekowitz, R.A.B.; Sastry, K.; Bailly, P.; Warner, A.
J. Exp. Med. 172, 1785-1794, 1990
A;Title: Molecular characterization of the human macrophage mannose receptor: demonstrat
A;Reference number: A60926; MUID:91079783; PMID:2258707
A;Accession: A60926
A;Status: nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1-1333,'T',1335-1456 <EZE>
A;Cross-references: GB:X55635
A;Note: translation of the nucleotide sequence is incomplete
A;Note: in the authors' translation additional residues Pro-Glu-Ile are shown after 497-
R;Kim, S.J.; Ruiz, N.; Bezouska, K.; Drickamer, K.
Genomics 14, 721-727, 1992
A;Title: Organization of the gene encoding the human macrophage mannose receptor (MRC1).
A;Reference number: A44255; MUID:93052405; PMID:1294118
A;Accession: A44255
A;Status: preliminary; nucleic acid sequence not shown; not compared with conceptual tra
A;Molecule type: DNA
A;Residues: 155-233,'KSAL',238-283;346-428;492-569;631-714,716-719;783-820,'N',822-865;'
A;Note: sequence extracted from NCBI backbone (NCBIP:118415, NCBIP:118421, NCBIP:118428,
C;Genetics:
A;Gene: GDB:MRC1
A;Cross-references: GDB:133759; OMIM:153618
A;Map position: 10p13-10p13
C;Superfamily: phospholipase A2 receptor; C-type lectin homology; fibronectin type II re
C;Keywords: duplication; lectin; tandem repeat; transmembrane protein
F;1-18/Domain: signal sequence #status predicted <SIG>
F;168-209/Domain: fibronectin type II repeat homology <2F1>
F;223-340/Domain: C-type lectin homology <LCH1>
F;362-486/Domain: C-type lectin homology <LCH2>
F;945-1079/Domain: C-type lectin homology <LCH3>

Query Match 9.2%; Score 183; DB 1; Length 1456;
Best Local Similarity 25.1%; Pred. No. 1.2e-06;
Matches 53; Conservative 38; Mismatches 66; Indels 54; Gaps 8;

Qy 42 YKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFWIGLR 101

```
                ||   :: :  :   : |:  |:|: | ||||:|| |:| : |:: |   :   ::|||
Db        807 YKDYQYYFSKEKETMDNARAFCKRNFGDLVSIQSESEKKFLWKYV-NRNDAQSAYFIGLL 865

Qy        102 RREEKQSNSTACQDLYAWTDGSISQFRNWYVDEPSCGS--EVCVVMYHQPSAPAGIGGPY 159
                :|:         :|| |||   : :|   ||:  :  | || ||    : :|
Db        866 ISLDKK---------FAWMDGSKVDYVSWATGEPNFANEDENCVTMY----SNSGF---- 908

Qy        160 MFQWNDDRCNMKNNFICKYSDEK----PAVPSREAEGEETELTTPVLPEETQE------- 208
                |||  |   | |||:  :        :|           | | :|   :|
Db        909 ---WNDINCGYPNAFICQRHNSSINATTVMP----------TMPSVPSGCKEGWNFYSN 954

Qy        209 ---------EDAKKTFKESREAALNLAYILI 230
                      |: :| ::|:|:| :     |:
Db        955 KCFKIFGFMEEERKNWQEARKACIGFGGNLV 985
```

RESULT 2
A48925
mannose receptor precursor, macrophage - mouse
C;Species: Mus musculus (house mouse)
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C;Accession: A48925; S21320; PC2245
R;Harris, N.; Super, M.; Rits, M.; Chang, G.; Ezekowitz, R.A.
Blood 80, 2363-2373, 1992
A;Title: Characterization of the murine macrophage mannose receptor: demonstration that
A;Reference number: A48925; MUID:93043353; PMID:1421407
A;Accession: A48925
A;Status: not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-1455 <HAR>
A;Experimental source: peritoneal macrophage
A;Note: sequence extracted from NCBI backbone (NCBIP:118733)
R;Harris, N.; Super, M.; Rits, M.; Chang, G.; Ezekowitz, R.B.
submitted to the EMBL Data Library, April 1992
A;Description: Characterization of the murine macrophage mannose receptor: Demonstration
on.
A;Reference number: S21320
A;Accession: S21320
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-302,'W',303-1117,'E',1119-1455 <HA2>
A;Cross-references: EMBL:Z11974; NID:g52997; PIDN:CAA78028.1; PID:g52998
R;Harris, N.; Peters, L.L.; Eicher, E.M.; Rits, M.; Raspberry, D.; Eichbaum, Q.G.; Super
Biochem. Biophys. Res. Commun. 198, 682-692, 1994
A;Title: The exon-intron structure and chromosomal localization of the mouse macrophage
A;Reference number: PC2245; MUID:94128116; PMID:8297379
A;Accession: PC2245
A;Molecule type: mRNA
A;Residues: 35-105 <HA3>
C;Genetics:
A;Gene: Mrc1
A;Map position: 2
C;Superfamily: phospholipase A2 receptor; C-type lectin homology; fibronectin type II re
C;Keywords: membrane protein; receptor
F;168-209/Domain: fibronectin type II repeat homology <2F9>
F;361-485/Domain: C-type lectin homology <LCH1>
F;943-1077/Domain: C-type lectin homology <LCH2>

  Query Match   8.9%; Score 178.5; DB 1; Length 1455;
  Best Local Similarity   25.9%; Pred. No. 2.7e-06;
  Matches   53; Conservative   33; Mismatches   66; Indels   53; Gaps   8;

```
Qy         42 YKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFWIGLR 101
              ||   :: :  :   : |:  |::: | | :|:|| |:| : |:| |       ::||:
Db        806 YKDYQYYFSKEKETMDNARRFCKKNFGDLATIKSESEKKFLWKYI-NKNGGQSPYFIGML 864

Qy        102 RREEKQSNSTACQDLYAWTDGSISQFRNWYVDEPSCGS--EVCVVMYHQPSAPAGIGGPY 159
                :|:         : | |||   |  |   ||:  :  | || ||      :|
Db        865 ISMDKK---------FIWMDGSKVDFVAWATGEPNFANDDENCVTMY----TNSGF---- 907

Qy        160 MFQWNDDRCNMKNNFICK---YSDEKPAVPSREAEGEETELTTPVLPEETQE-------- 208
                |||  |   |||||:    |    |:|           |||  |   :|
Db        908 ---WNDINCGYPNNFICQRHNSSINATAMP----------TTPTTPGGCKEGWHLYKNK 953
```

```
Qy        209 --------EDAKKTFKESREAALNL 225
                      : ||:::::|:|   |
Db        954 CFKIFGFANEEKKSWQDARQACKGL 978
```

RESULT 3
S52781
neurocan - mouse
C;Species: Mus musculus (house mouse)
C;Date: 19-May-1995 #sequence_revision 21-Jul-1995 #text_change 04-Feb-2000
C;Accession: S52781
R;Rauch, U.; Forsberg, N.; Kulbe, G.; Arnold-Ammer, I.; Faessler, R.
submitted to the EMBL Data Library, February 1995
A;Description: Amino acid sequence of mouse neurocan and brevican and their different exp
A;Reference number: S52781
A;Accession: S52781
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-1268 <RAU>
A;Cross-references: EMBL:X84727; NID:g758629; PIDN:CAA59216.1; PID:g758630
C;Superfamily: aggrecan; C-type lectin homology; complement factor H repeat homology; EGF
F;176-253/Domain: link protein repeat homology <LNK1>
F;274-355/Domain: link protein repeat homology <LNK2>
F;964-995/Domain: EGF homology <EGF>
F;1040-1160/Domain: C-type lectin homology <LCH>
F;1167-1223/Domain: complement factor H repeat homology <FHD>

  Query Match   8.8%; Score 177; DB 2; Length 1268;
  Best Local Similarity   31.5%; Pred. No. 3.1e-06;
  Matches   45; Conservative   17; Mismatches   49; Indels   32; Gaps   6;

```
Qy         38 QRPCYKVIYFHDTSRRLNPEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFW 97
              |  ||:  ||   : |  :|:|:  |||  | | |: | :| | |  |          : |
Db       1048 QGHCYR--YF---AHRRAWEDAERDCRRRAGHLTSVHSPEEHKFINSF------GHENSW 1096

Qy         98 IGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEPS---CGSEVCVVMYHQPSAPAG 154
              |||  |  ::         : ||| :  |: ||   :|      | | ||||    |
Db       1097 IGLNDRTVERD--------FQWTDNTGLQYENWREKQPDNFFAGGEDCVVMVAHESG--- 1145

Qy        155 IGGPYMFQWNDDRCNMKNNFICK 177
                :|||  ||    ::||
Db       1146 -------RWNDVPCNYNLPYVCK 1161
```

RESULT 4
T14274
versican precursor, splice form V2 - bovine
C;Species: Bos primigenius taurus (cattle)
C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 05-May-2000
C;Accession: T14274
R;Schmalfeldt, M.; Dours-Zimmermann, M.T.; Winterhalter, K.H.; Zimmermann, D.R.
J. Biol. Chem. 273, 15758-15764, 1998
A;Title: Versican V2 is a major extracellular matrix component of the mature bovine brain
A;Reference number: Z17954; MUID:98288320; PMID:9624174
A;Accession: T14274
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-1643 <SCH>
A;Cross-references: EMBL:AF060458; NID:g3253303; PID:g3253304; PIDN:AAC24360.1
A;Experimental source: brain
C;Keywords: glycoprotein
F;1-20/Domain: signal sequence #status predicted <SIG>
F;21-1643/Product: versican, splice form V2 #status predicted <MAT>
F;57,331,352,817,965,1017,1333,1616,1626/Binding site: carbohydrate (Asn) (covalent) #sta

  Query Match   8.7%; Score 174.5; DB 2; Length 1643;
  Best Local Similarity   25.5%; Pred. No. 6.8e-06;
  Matches   59; Conservative   26; Mismatches   71; Indels   75; Gaps   11;

```
Qy         38 QRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDF- 96
              |  |||  ||   : |  :: |:  ||  |  | || | :||  : :          |:
Db       1424 QGQCYK--YF---AHRRTWDAAERECRLQGAHLTSILSHEEQMFVNRV-------GHDYQ 1471
```

```
Qy          97 WIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEP----SCGSEVCVVMYHQPSAP 152
               ||||         |    :  : |||||  |: ||  ::|     | | :  |:::|:
Db        1472 WIGL--------NDKMFEHDFRWTDGSTLQYENWRPNQPDSFFSTGEDCVVIIWHENG-- 1521

Qy         153 AGIGGPYMFQWNDDRCNMKNNFICKYS----DEKPAVPSREAEGE--------------- 193
                   ||||  ||    : ||       : | | : :  |:
Db        1522 ---------QWNDVPCNYHLTYTCKKGTVACGQPPVVENAKTFGKMKPRYEINSLIRYHC 1572

Qy         194 -----ETELTT----------PVL----PEETQEEDAKKTFKESREAALN 224
                  :  | |           | :    |   |   :|| || |  |  |
Db        1573 KDGFIQRHLPTIRCLGNGRWAMPKITCLNPSAYQRTYSKKYFKNSSSAKDN 1623
```

RESULT 5
T42389
versican precursor, splice form V0 - bovine
N;Alternate names: chondroitin sulfate proteoglycan
C;Species: Bos primigenius taurus (cattle)
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 05-May-2000
C;Accession: T42389
R;Schmalfeldt, M.; Dours-Zimmermann, M.T.; Winterhalter, K.H.; Zimmermann, D.R.
J. Biol. Chem. 273, 15758-15764, 1998
A;Title: Versican V2 is a major extracellular matrix component of the mature bovine brai
A;Reference number: Z17954; MUID:98288320; PMID:9624174
A;Accession: T42389
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-3381 <SCH>
A;Cross-references: EMBL:AF060456; NID:g3253299; PID:g3253300; PIDN:AAC24358.1
C;Superfamily: chicken chondroitin sulfate proteoglycan PG-M core protein; C-type lectin
C;Keywords: chondroitin sulfate proteoglycan; extracellular matrix; glycoprotein
F;1-20/Domain: signal sequence #status predicted <SIG>
F;21-3381/Product: versican, splice form V0 #status predicted <MAT>
F;57,331,352,817,965,1017,1333,1393,1437,1463,1653,1974,2045,2074,2103,2263,2290,2356,26

  Query Match   8.7%; Score 174.5; DB 2; Length 3381;
  Best Local Similarity   25.5%; Pred. No. 1.7e-05;
  Matches   59; Conservative   26; Mismatches   71; Indels   75; Gaps   11;

```
Qy          38 QRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDF- 96
               |  |||  ||   : |  :: |:  ||  |  | || | :||  : :          |:
Db        3162 QGQCYK--YF---AHRRTWDAAERECRLQGAHLTSILSHEEQMFVNRV-------GHDYQ 3209

Qy          97 WIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEP----SCGSEVCVVMYHQPSAP 152
               ||||         |    :  : |||||  |: ||  ::|     | | :  |:::|:
Db        3210 WIGL--------NDKMFEHDFRWTDGSTLQYENWRPNQPDSFFSTGEDCVVIIWHENG-- 3259

Qy         153 AGIGGPYMFQWNDDRCNMKNNFICKYS----DEKPAVPSREAEGE--------------- 193
                   ||||  ||    : ||       : | | : :  |:
Db        3260 ---------QWNDVPCNYHLTYTCKKGTVACGQPPVVENAKTFGKMKPRYEINSLIRYHC 3310

Qy         194 -----ETELTT----------PVL----PEETQEEDAKKTFKESREAALN 224
                  :  | |           | :    |   |   :|| || |  |  |
Db        3311 KDGFIQRHLPTIRCLGNGRWAMPKITCLNPSAYQRTYSKKYFKNSSSAKDN 3361
```

RESULT 6
S28764
neurocan precursor - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 22-Nov-1993 #sequence_revision 01-Sep-1995 #text_change 04-Feb-2000
C;Accession: S28764
R;Rauch, U.; Karthikeyan, L.; Maurel, P.; Margolis, R.U.; Margolis, R.K.
J. Biol. Chem. 267, 19536-19547, 1992
A;Title: Cloning and primary structure of neurocan, a developmentally regulated, aggrega
A;Reference number: S28764; MUID:92406907; PMID:1326557
A;Accession: S28764
A;Molecule type: mRNA
A;Residues: 1-1257 <RAU>
A;Cross-references: EMBL:M97161; NID:g205649; PIDN:AAC37679.1; PID:g205650
C;Superfamily: aggrecan; C-type lectin homology; complement factor H repeat homology; EG
C;Keywords: chondroitin sulfate proteoglycan; glycoprotein
F;1-22/Domain: signal sequence #status predicted <SIG>
F;23-1257/Product: neurocan #status predicted <MAT>
F;176-253/Domain: link protein repeat homology <LNK1>
F;274-355/Domain: link protein repeat homology <LNK2>
F;364-366/Region: cell attachment (R-G-D) motif
F;953-984/Domain: EGF homology <EGF>
F;1029-1149/Domain: C-type lectin homology <LCH>
F;1156-1212/Domain: complement factor H repeat homology <FHD>
F;121,339,737,967,1164/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;372,410/Binding site: chondroitin sulfate (Ser) (covalent) #status predicted
F;944/Binding site: chondroitin sulfate (Ser) (covalent) #status experimental

  Query Match   8.7%; Score 174; DB 2; Length 1257;
  Best Local Similarity   30.8%; Pred. No. 5.4e-06;
  Matches   44; Conservative   18; Mismatches   49; Indels   32; Gaps   6;

```
Qy          38 QRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFW 97
               |  ||:  ||   : |  :|:|:  |||  | | |: | :| | |  |         : |
Db        1037 QGHCYR--YF---AHRRAWEDAERDCRRRAGHLTSVHSPEEHKFINSF------GHENSW 1085

Qy          98 IGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEPS---CGSEVCVVMYHQPSAPAG 154
               |||  |  ::          : ||| :  |: ||   :|     | | ||||    :
Db        1086 IGLNDRTVERD--------FQWTDNTGLQYENWREKQPDNFFAGGEDCVVMVAHENG--- 1134

Qy         155 IGGPYMFQWNDDRCNMKNNFICK 177
                  :|||  ||    ::||
Db        1135 -------RWNDVPCNYNLPYVCK 1150
```

RESULT 7
A55535
versican precursor - mouse
N;Alternate names: chondroitin sulfate proteoglycan 2; chondroitin sulfate proteoglycan versican
N;Contains: glial hyaluronate-binding protein
C;Species: Mus musculus (house mouse)
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C;Accession: A55535
R;Ito, K.; Shinomura, T.; Zako, M.; Ujita, M.; Kimata, K.
J. Biol. Chem. 270, 958-965, 1995
A;Title: Multiple forms of mouse PG-M, a large chondroitin sulfate proteoglycan generate
A;Reference number: A55535; MUID:95122551; PMID:7822336
A;Accession: A55535
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-2397 <RES>
A;Cross-references: GB:D16263; NID:g862460; PIDN:BAA03796.1; PID:g862461
C;Superfamily: versican; C-type lectin homology; complement factor H repeat homology; EG
F;1-20/Domain: signal sequence #status predicted <SIG>
F;21-1654/Domain: versican #status predicted <MAT>
F;167-244/Domain: link protein repeat homology <LNK1>
F;265-346/Domain: link protein repeat homology <LNK2>
F;2095-2126/Domain: EGF homology <EG1>
F;2133-2164/Domain: EGF homology <EG2>
F;2171-2291/Domain: C-type lectin homology <LCH>
F;2298-2354/Domain: complement factor H repeat homology <FHD>

  Query Match   8.7%; Score 174; DB 1; Length 2397;
  Best Local Similarity   28.5%; Pred. No. 1.2e-05;
  Matches   47; Conservative   23; Mismatches   55; Indels   40; Gaps   8;

```
Qy          38 QRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDF- 96
               |  |||  ||   : |  :: |:  ||  |  | || | :||  : :          |:
Db        2179 QGQCYK--YF---AHRRTWDAAERECRLQGAHLTSILSHEEQMFVNRV-------GHDYQ 2226

Qy          97 WIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEP----SCGSEVCVVMYHQPSAP 152
               ||||         |    :  : |||||  |: ||  ::|     | | :  |:::|:
Db        2227 WIGL--------NDKMFEHDFRWTDGSALQYENWRPNQPDSFFSAGEDCVVIIWHENG-- 2276

Qy         153 AGIGGPYMFQWNDDRCNMKNNFICKYS----DEKPAVPSREAEGE 193
                   ||||  ||    : ||       : | | : :  |:
Db        2277 ---------QWNDVPCNYHLTYTCKKGTVACGQPPVVENAKTFGK 2312
```

RESULT 8
A60979
versican precursor - human
N;Alternate names: chondroitin sulfate proteoglycan 2; chondroitin sulfate proteoglycan
N;Contains: glial hyaluronate-binding protein
C;Species: Homo sapiens (man)
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 19-Jan-2001
C;Accession: S06014; S43921; A60979; A30358; A29348; A45131; I54179
R;Zimmermann, D.R.; Ruoslahti, E.
EMBO J. 8, 2975-2981, 1989
A;Title: Multiple domains of the large fibroblast proteoglycan, versican.
A;Reference number: S06014; MUID:90059882; PMID:2583089
A;Accession: S06014
A;Molecule type: mRNA
A;Residues: 1-2409 <ZIM>
A;Cross-references: GB:X15998; NID:g37662; PIDN:CAA34128.1; PID:g37663
R;Yao, L.Y.; Moody, C.; Schoenherr, E.; Wight, T.N.; Sandell, L.J.
Matrix Biol. 14, 213-225, 1994
A;Title: Identification of the proteoglycan versican in aorta and smooth muscle cells by
A;Reference number: S43921; MUID:95005762; PMID:7921538
A;Accession: S43921
A;Molecule type: mRNA
A;Residues: 208-440;1094-1385;1910-2246 <YAO>
R;Bignami, A.; Lane, W.S.; Andrews, D.; Dahl, D.
Brain Res. Bull. 22, 67-70, 1989
A;Title: Structural similarity of hyaluronate binding proteins in brain and cartilage.
A;Reference number: A60979; MUID:89229983; PMID:2469524
A;Accession: A60979
A;Molecule type: protein
A;Residues: 171-210;289-303 <BIG>
R;Perides, G.; Lane, W.S.; Andrews, D.; Dahl, D.; Bignami, A.
J. Biol. Chem. 264, 5981-5987, 1989
A;Title: Isolation and partial characterization of a glial hyaluronate-binding protein.
A;Reference number: A30358; MUID:89174663; PMID:2466833
A;Accession: A30358
A;Molecule type: protein
A;Residues: 24-50;80-87,'D',89-119;128-155;167-218;229-259,'IR';261-268;277-283,'G',285-
R;Krusius, T.; Gehlsen, K.R.; Ruoslahti, E.
J. Biol. Chem. 262, 13120-13125, 1987
A;Title: A fibroblast chondroitin sulfate proteoglycan core protein contains lectin-like
A;Reference number: A29348; MUID:88007514; PMID:2820964
A;Accession: A29348
A;Molecule type: mRNA
A;Residues: 1725,'V',1727-2409 <KRU>
A;Cross-references: GB:J02814
R;Perides, G.; Rahemtulla, F.; Lane, W.S.; Asher, R.A.; Bignami, A.
J. Biol. Chem. 267, 23883-23887, 1992
A;Title: Isolation of a large aggregating proteoglycan from human brain.
A;Reference number: A45131; MUID:93054750; PMID:1429726
A;Contents: brain
A;Accession: A45131
A;Molecule type: protein
A;Residues: 21-22,'X',24-37 <PE2>
A;Experimental source: brain
A;Note: sequence extracted from NCBI backbone (NCBIP:118884)
R;Iozzo, R.V.; Naso, M.F.; Cannizzaro, L.A.; Wasmuth, J.J.; McPherson, J.D.
Genomics 14, 845-851, 1992
A;Title: Mapping of the versican proteoglycan gene (CSPG2) to the long arm of human chro
A;Reference number: I54179; MUID:93122792; PMID:1478664
A;Accession: I54179
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 251-347 <RES>
A;Cross-references: GB:S52488; NID:g263313; PIDN:AAB24878.1; PID:g263314
C;Genetics:
A;Gene: GDB:CSPG2
A;Cross-references: GDB:127873; OMIM:118661
A;Map position: 5q12-5q14
C;Superfamily: versican; C-type lectin homology; complement factor H repeat homology; EG
F;1-20/Domain: signal sequence #status predicted <SIG>
F;21-2409/Product: proteoglycan 24K core protein #status predicted <MAT>
F;167-244/Domain: link protein repeat homology <LNK1>
F;265-346/Domain: link protein repeat homology <LNK2>
F;559-1654/Domain: chondroitin sulfate attachment #status predicted <GAG>
F;2106-2137/Domain: EGF homology <EG1>
F;2144-2175/Domain: EGF homology <EG2>
F;2182-2302/Domain: C-type lectin homology <LCH>
F;2309-2365/Domain: complement factor H repeat homology <FHD>

Query Match 8.7%; Score 174; DB 1; Length 2409;
Best Local Similarity 28.5%; Pred. No. 1.2e-05;
Matches 47; Conservative 23; Mismatches 55; Indels 40; Gaps 8;

```
Qy        38 QRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDF- 96
             |  |||  ||   : |  :: |:  ||  |  | || | :||  : :          |:
Db      2190 QGQCYK--YF---AHRRTWDAAERECRLQGAHLTSILSHEEQMFVNRV-------GHDYQ 2237

Qy        97 WIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEP----SCGSEVCVVMYHQPSAP 152
             ||||        |    :  : |||||  |: ||  ::|    | | :  |:::|:
Db      2238 WIGL--------NDKMFEHDFRWTDGSTLQYENWRPNQPDSFFSAGEDCVVIIWHENG-- 2287

Qy       153 AGIGGPYMFQWNDDRCNMKNNFICKYS----DEKPAVPSREAEGE 193
                    ||||  ||    : ||       : | | : :  |:
Db      2288 ---------QWNDVPCNYHLTYTCKKGTVACGQPPVVENAKTFGK 2323
```

RESULT 9
A47171
chondroitin sulfate proteoglycan PG-M core protein - chicken
C;Species: Gallus gallus (chicken)
C;Date: 21-Sep-1993 #sequence_revision 18-Nov-1994 #text_change 21-Jul-2000
C;Accession: A47171
R;Shinomura, T.; Nishida, Y.; Ito, K.; Kimata, K.
J. Biol. Chem. 268, 14461-14469, 1993
A;Title: cDNA cloning of PG-M, a large chondroitin sulfate proteoglycan expressed during
A;Reference number: A47171; MUID:93300846; PMID:8314802
A;Accession: A47171
A;Status: preliminary
A;Molecule type: nucleic acid
A;Residues: 1-3562 <SHI>
A;Cross-references: GB:D13542; NID:g391643; PIDN:BAA02742.1; PID:g391644
A;Experimental source: stage 22-23 developing limb buds
A;Note: sequence extracted from NCBI backbone (NCBIN:134456, NCBIP:134457)
C;Superfamily: chicken chondroitin sulfate proteoglycan PG-M core protein; C-type lectin
F;166-243/Domain: link protein repeat homology <LNK1>
F;264-345/Domain: link protein repeat homology <LNK2>
F;3258-3289/Domain: EGF homology <EGF1>
F;3296-3327/Domain: EGF homology <EGF>
F;3334-3454/Domain: C-type lectin homology <LCH>
F;3461-3517/Domain: complement factor H repeat homology <FHD>

Query Match 8.6%; Score 171; DB 2; Length 3562;
Best Local Similarity 28.5%; Pred. No. 3.5e-05;
Matches 47; Conservative 23; Mismatches 55; Indels 40; Gaps 8;

```
Qy        38 QRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDF- 96
             |  |||  ||   : |  :: |:  ||  |  | || | :||  : :          |:
Db      3342 QGQCYK--YF---AHRRTWDTAERECRLQGAHLTSILSHEEQVFVNRI-------GHDYQ 3389

Qy        97 WIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEP----SCGSEVCVVMYHQPSAP 152
             ||||        |    :  : |||||  |: ||  ::|    | | :  |:::|:
Db      3390 WIGL--------NDKMFERDFRWTDGSPLQYENWRPNQPDSFFSAGEDCVVIIWHENG-- 3439

Qy       153 AGIGGPYMFQWNDDRCNMKNNFICKYS----DEKPAVPSREAEGE 193
                    ||||  ||    : ||       : | | : :  |:
Db      3440 ---------QWNDVPCNYHLTYTCKKGTVACGQPPVVENAKTFGK 3475
```

RESULT 10
T42710
mannose receptor, macrophage - mouse
N;Alternate names: lambda lectin; phospholipase A2 receptor
C;Species: Mus musculus (house mouse)
C;Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 09-Jun-2000

```
C;Accession: T42710
R;Wu, K.; Yuan, J.; Lasky, L.A.
J. Biol. Chem. 271, 21323-21330, 1996
A;Title: Characterization of a novel member of the macrophage mannose receptor type C le
A;Reference number: Z22235; MUID:96355501; PMID:8702911
A;Accession: T42710
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-1479 <WUK>
A;Cross-references: EMBL:U56734; NID:g1336073; PID:g1336074; PIDN:AAC52729.1
C;Superfamily: phospholipase A2 receptor; C-type lectin homology; fibronectin type II re
C;Keywords: membrane protein; receptor
F;186-227/Domain: fibronectin type II repeat homology <2FR>

  Query Match             8.5%;  Score 170.5;  DB 2;  Length 1479;
  Best Local Similarity   31.4%;  Pred. No. 1.3e-05;
  Matches   48;  Conservative   22;  Mismatches   48;  Indels   35;  Gaps    7;

Qy         28 SGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIE 87
              | ||      |  ||::      : : :::|:| || | || |:|| |  | : | | |: 
Db        384 SWQPF-----QGHCYRL-----QAEKRSWQESKRACLRGGGDLLSIHSMAELEFITKQIK 433

Qy         88 NLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEPS---CGSEVCVV 144
                :    : ||||          |     |  : |:|||:  | :|:  ||:       | ||
Db        434 QEVE---ELWIGL--------NDLKLQMNFEWSDGSLVSFTHWHPFEPNNFRDSLEDCVT 482

Qy        145 MYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICK 177
              ::     | |       :|||  ||     |||
Db        483 IW----GPEG-------RWNDSPCNQSLPSICK 504


RESULT 11
A39808
proteoglycan core protein, cartilage - bovine (fragments)
N;Alternate names: aggrecan; aggregating cartilage proteoglycan
C;Species: Bos primigenius taurus (cattle)
C;Date: 20-Mar-1992 #sequence_revision 23-Mar-1995 #text_change 13-Aug-1999
C;Accession: A34234; A27752; A39808; A27751; E29164; B27751; C27751; D27751; E27751; F27
R;Antonsson, P.; Heinegard, D.; Oldberg, A.
J. Biol. Chem. 264, 16170-16173, 1989
A;Title: The keratan sulfate-enriched region of bovine cartilage proteoglycan consists c
A;Reference number: A34234; MUID:89380219; PMID:2528543
A;Accession: A34234
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 128-621 <ANT>
A;Cross-references: GB:J05028
R;Oldberg, A.; Antonsson, P.; Heinegard, D.
Biochem. J. 243, 255-259, 1987
A;Title: The partial amino acid sequence of bovine cartilage proteoglycan,deduced from a
A;Reference number: A27752; MUID:87270630; PMID:3111460
A;Accession: A27752
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 622-1340 <OLD>
R;Sandy, J.D.; Boynton, R.E.; Flannery, C.R.
J. Biol. Chem. 266, 8198-8205, 1991
A;Title: Analysis of the catabolism of aggrecan in cartilage explants by quantitation of
A;Reference number: A39808; MUID:91217051; PMID:2022637
A;Accession: A39808
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-28;59-82;131-137,'QSET',142-149;196-207;226-249;1137-1143;1252-1267;1274-1
R;Perin, J.P.; Bonnet, F.; Jolles, P.
FEBS Lett. 206, 73-77, 1986
A;Title: Structural relationship between link proteins and proteoglycan monomers.
A;Reference number: A27751; MUID:87005253; PMID:3530809
A;Accession: A27751
A;Molecule type: protein
A;Residues: 29-58;74-130;174-175,'A',177-204;208-225 <PER>
R;Perin, J.P.; Bonnet, F.; Jolles, J.; Jolles, P.
FEBS Lett. 176, 37-42, 1984
A;Title: Sequence data concerning the protein core of the cartilage proteoglycan monomer
A;Reference number: A91327; MUID:85027710; PMID:6489519
A;Accession: E29164
A;Molecule type: protein
A;Residues: 1230-1249 <PE2>
C;Superfamily: aggrecan; C-type lectin homology; complement factor H repeat homology; EGI
C;Keywords: glycoprotein
F;8-28/Domain: link protein repeat homology (fragment) <LNK1>
F;29-58/Domain: link protein repeat homology (fragment) <LNK2>
F;80-146/Domain: link protein repeat homology (fragments) <LNK3>
F;167-248/Domain: link protein repeat homology <LNK4>
F;1130-1250/Domain: C-type lectin homology <LCH>
F;1257-1313/Domain: complement factor H repeat homology <FHD>

  Query Match             7.9%;  Score 158.5;  DB 2;  Length 1340;
  Best Local Similarity   26.8%;  Pred. No. 0.00011;
  Matches   48;  Conservative   29;  Mismatches   59;  Indels   43;  Gaps   10;

Qy         30 QPVCRGG---TQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFI 86
              | :|  |     |  ||:  :| |   |  : :|:  ||:     | || : :||    :|:
Db       1127 QKLCEEGWTKFQGHCYR--HFPD---RATWVDAESQCRKQQSHLSSIVTPEEQ----EFV 1177

Qy         87 ENLLPSDGDF-WIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEP----SCGSEV 141
                |   :  |: ||||         |     :  : |:||    || ||  ::|     : | :
Db       1178 NN---NAQDYQWIGL--------NDKTIEGDFRWSDGHSLQFENWRPNQPDNFFATGEDC 1226

Qy        142 CVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYS----DEKPAVPSREAEGEETE 196
               |:::|:             :|||  || :  | ||       | | |      |:: :
Db       1227 VVMIWHEKG----------EWNDVPCNYQLPFTCKKGTVACGEPPVVEHARIFGQKKD 1274


RESULT 12
T42630
aggrecan - bovine
C;Species: Bos primigenius taurus (cattle)
C;Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 04-Mar-2000
C;Accession: T42630
R;Hering, T.M.; Kollar, J.; Huynh, T.D.
submitted to the EMBL Data Library, September 1996
A;Description: Complete coding sequence of bovine aggrecan: comparative structural analy
A;Reference number: Z22182
A;Accession: T42630
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-2327 <HER>
A;Cross-references: EMBL:U76615; NID:g1730259; PID:g1730260; PIDN:AAB38524.1
A;Experimental source: articular chondrocytes
C;Superfamily: aggrecan; C-type lectin homology; complement factor H repeat homology; EG
C;Keywords: cartilage; chondroitin sulfate proteoglycan; extracellular matrix; glycoprot

  Query Match             7.9%;  Score 158.5;  DB 2;  Length 2327;
  Best Local Similarity   26.8%;  Pred. No. 0.00022;
  Matches   48;  Conservative   29;  Mismatches   59;  Indels   43;  Gaps   10;

Qy         30 QPVCRGG---TQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFI 86
              | :|  |     |  ||:  :| |   |  : :|:  ||:     | || : :||    :|:
Db       2114 QKLCEEGWTKFQGHCYR--HFPD---RATWVDAESQCRKQQSHLSSIVTPEEQ----EFV 2164

Qy         87 ENLLPSDGDF-WIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEP----SCGSEV 141
                |   :  |: ||||         |     :  : |:||    || ||  ::|     : | :
Db       2165 NN---NAQDYQWIGL--------NDKTIEGDFRWSDGHSLQFENWRPNQPDNFFATGEDC 2213

Qy        142 CVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYS----DEKPAVPSREAEGEETE 196
               |:::|:             :|||  || :  | ||       | | |      |:: :
Db       2214 VVMIWHEKG----------EWNDVPCNYQLPFTCKKGTVACGEPPVVEHARIFGQKKD 2261


RESULT 13
A39086
aggrecan precursor, cartilage long splice form [validated] - human
N;Alternate names: chondroitin sulfate proteoglycan 1; large aggregating proteoglycan; p
N;Contains: aggrecan cartilage short splice form
C;Species: Homo sapiens (man)
C;Date: 10-Sep-1999 #sequence_revision 01-Dec-2000 #text_change 08-Dec-2000
```

C;Accession: A39086; S50206; A43919; S46659; S66389; S68646; S62786; A34226; B43919; C43
R;Doege, K.J.; Sasaki, M.; Kimura, T.; Yamada, Y.
J. Biol. Chem. 266, 894-902, 1991
A;Title: Complete coding sequence and deduced primary structure of the human cartilage l
A;Reference number: A39086; MUID:91093289; PMID:1985970
A;Accession: A39086
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-2162,2201-2329,'A',2392-2415 <DOE>
A;Cross-references: GB:M55172; NID:g178258; PIDN:AAA62824.1; PID:g178259
R;Glumoff, V.; Savontaus, M.; Vehanen, J.; Vuorio, E.
Biochim. Biophys. Acta 1219, 613-622, 1994
A;Title: Analysis of aggrecan and tenascin gene expression in mouse skeletal tissues by
A;Reference number: S50206; MUID:95035091; PMID:7524681
A;Accession: S50206
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 350-497 <GLU>
A;Cross-references: EMBL:X80278; NID:g516295
A;Note: this translation is not annotated in GenBank entry HSAGGREC, release 113.0
R;Sandy, J.D.; Flannery, C.R.; Neame, P.J.; Lohmander, L.S.
J. Clin. Invest. 89, 1512-1516, 1992
A;Title: The structure of aggrecan fragments in human synovial fluid. Evidence for the i
domain.
A;Reference number: A43919; MUID:92235266; PMID:1569188
A;Accession: A43919
A;Molecule type: protein
A;Residues: 361-370,'X',372-373;393-399,'X',401-407,'X',409 <SAN>
A;Cross-references: PIDN:AAB22079.1; PID:g248844; PIDN:AAB22077.1; PID:g248842; PIDN:AAB
A;Experimental source: synovial fluid
A;Note: sequences modified after extraction from NCBI backbone
R;Barry, F.P.; Neame, P.J.; Sasse, J.; Pearson, D.
Matrix Biol. 14, 323-328, 1994
A;Title: Length variation in the keratan sulfate domain of mammalian aggrecan.
A;Reference number: I46998; MUID:95128522; PMID:7827755
A;Accession: S46659
A;Molecule type: DNA
A;Residues: 764-765,'A',767-846,'V',848-862,'X',864 <BAR>
A;Cross-references: EMBL:S74659; NID:g807127; PIDN:AAC60643.1; PID:g807128
A;Note: the authors translated the codon GAA for residue 803 as Ala and CYT for residue
R;Ilic, M.Z.; Mok, M.T.; Williamson, O.D.; Campbell, M.A.; Hughes, C.E.; Handley, C.J.
Arch. Biochem. Biophys. 322, 22-30, 1995
A;Title: Catabolism of aggrecan by explant cultures of human articular cartilage in the
A;Reference number: S66389; MUID:96004775; PMID:7574678
A;Accession: S66389
A;Status: preliminary
A;Molecule type: protein
A;Residues: 17-23;24,'X',26-27;393-401;402-403 <ILI>
R;Fosang, A.J.; Last, K.; Knaeuper, V.; Murphy, G.; Neame, P.J.
FEBS Lett. 380, 17-20, 1996
A;Title: Degradation of cartilage aggrecan by collagenase-3 (MMP-13).
A;Reference number: S68646; MUID:96181659; PMID:8603731
A;Accession: S68646
A;Status: preliminary
A;Molecule type: protein
A;Residues: 'V',404-405,'XX' <FOS>
R;Dudhia, J.; Davidson, C.M.; Wells, T.M.; Vynios, D.H.; Hardingham, T.E.; Bayliss, M.T.
Biochem. J. 313, 933-940, 1996
A;Title: Age-related changes in the content of the C-terminal region of aggrecan in huma
A;Reference number: S62786; MUID:96190740; PMID:8611178
A;Accession: S62786
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: mRNA
A;Residues: 1778-1927,'A',1929-1963,'V',1965-2162,2201-2415 <DUD>
A;Cross-references: EMBL:X17406; NID:g30248; PIDN:CAA35463.1; PID:g30249
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, January 1990
R;Baldwin, C.T.; Reginato, A.M.; Prockop, D.J.
J. Biol. Chem. 264, 15747-15750, 1989
A;Title: A new epidermal growth factor-like domain in the human core protein for the lar
A;Reference number: A34226; MUID:89380154; PMID:2789216
A;Accession: A34226
A;Molecule type: mRNA
A;Residues: 1936-1963,'V',1965-2069,'A',2071-2415 <BAL>
A;Cross-references: GB:J05062; NID:g181167; PIDN:AAA35726.1; PID:g181168
C;Genetics:
A;Gene: GDB:AGC1; CSPG1; CSPGCP; MSK16
A;Cross-references: GDB:127479; OMIM:155760
A;Map position: 15q26-15q26
C;Superfamily: aggrecan; C-type lectin homology; complement factor H repeat homology; EGF
C;Keywords: alternative splicing; cartilage; chondroitin sulfate proteoglycan; extracellu
F;1-19/Domain: signal sequence #status predicted <SIG>
F;20-2415/Product: aggrecan cartilage long splice form #status predicted <MATL>
F;20-2162,2201-2415/Product: aggrecan cartilage short splice form #status predicted <MATN
F;20-2162,2201-2329,'A',2392-2415/Product: aggrecan short splice form #status predicted <
F;44-135/Domain: immunoglobulin homology <IMM>
F;170-247/Domain: link protein repeat homology <LNK1>
F;268-349/Domain: link protein repeat homology <LNK2>
F;495-572/Domain: link protein repeat homology <LNK3>
F;593-673/Domain: link protein repeat homology <LNK4>
F;677-861/Domain: keratan sulfate attachment #status predicted <KSA>
F;864-1510/Domain: chondroitin sulfate attachment #status predicted <CS1>
F;1511-2162/Domain: chondroitin sulfate attachment #status predicted <CS2>
F;2168-2198/Domain: EGF homology <EGF>
F;2205-2325/Domain: C-type lectin homology <LCH>
F;2332-2388/Domain: complement factor H repeat homology <FHD>
F;126,239,333,387,434,602,657,737,1898/Binding site: carbohydrate (Asn) (covalent) #statu
F;371,376/Binding site: keratan sulfate (Thr) (covalent) #status predicted

Query Match 7.9%; Score 158.5; DB 1; Length 2415;
Best Local Similarity 29.1%; Pred. No. 0.00023;
Matches 52; Conservative 22; Mismatches 62; Indels 43; Gaps 11;

```
Qy         30 QPVCRGG---TQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFI 86
              | ||  |    |  ||:  :| |   |  : :|:  ||     | || : :||    :|:
Db       2202 QEVCEEGWNKYQGHCYR--HFPD---RETWVDAERRCREQQSHLSSIVTPEEQ----EFV 2252

Qy         87 ENLLPSDGDF-WIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEPS---CGSEVC 142
              |   :  |: ||||       |    :  : |:||   || ||  ::|       | |
Db       2253 NN---NAQDYQWIGL--------NDRTIEGDFRWSDGHPMQFENWRPNQPDNFFAAGEDC 2301

Qy        143 VVM-YHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYS----DEKPAVPSREAEGEETE 196
              ||| :|:             :|||  ||    | ||       | | |      |:: :
Db       2302 VVMIWHEKG----------EWNDVPCNYHLPFTCKKGTVACGEPPVVEHARTFGQKKD 2349
```

RESULT 14
LNRC1
lectin BRA3-1 precursor - barnacle (Megabalanus rosa)
C;Species: Megabalanus rosa
C;Date: 24-Feb-1994 #sequence_revision 09-Sep-1994 #text_change 16-Jul-1999
C;Accession: JC1503; A26094
R;Takamatsu, N.; Takeda, T.; Kojima, M.; Heishi, M.; Muramoto, K.; Kamiya, H.; Shiba, T.
Gene 128, 251-255, 1993
A;Title: Acorn barnacle Megabalanus rosa lectin (BRA-3): cDNA cloning, gene structure and
A;Reference number: JC1503; MUID:93292994; PMID:8514190
A;Accession: JC1503
A;Molecule type: DNA; mRNA
A;Residues: 1-162 <TAK>
A;Cross-references: DDBJ:D13299
R;Muramoto, K.; Kamiya, H.
Biochim. Biophys. Acta 874, 285-295, 1986
A;Title: The amino-acid sequence of a lectin of the acorn barnacle Megabalanus rosa.
A;Reference number: A26094
A;Accession: A26094
A;Molecule type: protein
A;Residues: 25-145,'K',147-162 <MUR>
A;Note: 146-Arg was also found
C;Comment: This galactose-binding lectin is isolated from the coelomic fluid.
C;Comment: This protein plays important roles in defense mechanisms and in development ar
C;Comment: The molecule is a tetramer of identical chains.
C;Genetics:
A;Introns: 22/1; 47/2; 86/3
C;Superfamily: tetranectin; C-type lectin homology
C;Keywords: hemolymph; homotetramer; lectin
F;1-24/Domain: signal sequence #status predicted <SIG>
F;25-162/Product: lectin BRA3-1 #status experimental <MAT>

```
F;26-150/Domain: C-type lectin homology <LCH>
F;26-39,56-150,125-142/Disulfide bonds: #status experimental
F;157/Disulfide bonds: interchain (to 160) #status experimental
F;160/Disulfide bonds: interchain (to 157) #status experimental

  Query Match             7.8%;  Score 156;  DB 1;  Length 162;
  Best Local Similarity   24.9%;  Pred. No. 1.3e-05;
  Matches   46;  Conservative   30;  Mismatches   67;  Indels   42;  Gaps    7;

Qy          1 MRPGTALQAVLLAVLLVGLRAATGRLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAK 60
              |:    :|||  | |::         : :|:  | |       |    :  :: :: : :|:
Db          1 MQRSEIVQAVTLLVVVFA-------ITTGECTCPGNLDWQEYDGHCYWASTYQVRWNDAQ 53

Qy         61 EACR--RDGGQLVSIESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYA 118
               ||:    |  | :|:|: |    | : : |         ||||         |    :  |
Db         54 LACQTVHPGAYLATIQSQLENAFISETVSN-----NRLWIGL--------NDIDLEGHYV 100

Qy        119 WTDGSISQFRNWYVDEPS------CGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKN 172
              |::|  :  |  |  : |:      ||    || |   :           ||:|| ||
Db        101 WSNGEATDFTYWSSNNPNNWENQDCG----VVNYDTVTG----------QWDDDDCNKNR 146

Qy        173 NFICK 177
              ||:||
Db        147 NFLCK 151


RESULT 15
B42755
E-selectin precursor - mouse
N;Alternate names: endothelial leukocyte adhesion molecule 1 (ELAM-1)
C;Species: Mus musculus (house mouse)
C;Date: 04-Mar-1993 #sequence_revision 31-Dec-2000 #text_change 02-Aug-2002
C;Accession: S23174; B42755
R;Becker-Andre, M.; van Huijsduijnen, R.H.; Losberger, C.; Whelan, J.; Delamarter, J.F.
Eur. J. Biochem. 206, 401-411, 1992
A;Title: Murine endothelial leukocyte-adhesion molecule 1 is a close structural and func
A;Reference number: S23174; MUID:92283265; PMID:1375914
A;Accession: S23174
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-612 <BEC>
A;Cross-references: GB:M80778; NID:g193014; PIDN:AAA37547.1; PID:g193015
R;Weller, A.; Isenmann, S.; Vestweber, D.
J. Biol. Chem. 267, 15176-15183, 1992
A;Title: Cloning of the mouse endothelial selectins. Expression of both E- and P-selecti
A;Reference number: A42755; MUID:92340571; PMID:1378846
A;Accession: B42755
A;Status: nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 'MKATAGV',1-389,391-612 <WEL>
A;Cross-references: GB:M87862; NID:g193107
A;Experimental source: endothelial cells
A;Note: sequence extracted from NCBI backbone (NCBIP:109470)
A;Note: the sequence in GenBank entry MUSESELEC, release 117.0, (PIDN:AAA37577.1; PID:g1
A;Note: it is uncertain whether the initiator is Met-1 or the AUG codon preceding that
C;Superfamily: P-selectin; C-type lectin homology; complement factor H repeat homology;
C;Keywords: glycoprotein; transmembrane protein
F;1-21/Domain: signal sequence #status predicted <SIG>
F;12-138/Domain: C-type lectin homology <LCH>
F;22-612/Product: P-selectin #status predicted <MAT>
F;143-174/Domain: EGF homology <EGF>
F;180-238/Domain: complement factor H repeat homology <FH1>
F;243-300/Domain: complement factor H repeat homology <FH2>
F;305-363/Domain: complement factor H repeat homology <FH3>
F;368-426/Domain: complement factor H repeat homology <FH4>
F;431-489/Domain: complement factor H repeat homology <FH5>
F;494-548/Domain: complement factor H repeat homology <FH6>
F;25,391,528/Binding site: carbohydrate (Asn) (covalent) #status predicted

  Query Match             7.7%;  Score 154.5;  DB 2;  Length 612;
  Best Local Similarity   27.9%;  Pred. No. 9e-05;
  Matches   38;  Conservative   28;  Mismatches   45;  Indels   25;  Gaps    5;

Qy         46 YFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFWIGLRRREE 105
              |:: :|  : ::||   |:||   ||:|::::|   |     ||  |   :|||:|:
Db         23 YYNASSELMTYDEASAYCQRDYTHLVAIQNKEE---INYLNSNLKHSPSYYWIGIRK--- 76

Qy        106 KQSNSTACQDLYAWTDGS---ISQFRNWYVDEPS--CGSEVCVVMYHQPSAPAGIGGPYM 160
                     ::: |       : :||   ||:    :| || :| | :  :|:
Db         77 -------VNNVWIWVGTGKPLTEEAQNWAPGEPNNKQRNEDCVEIYIQRTKDSGM----- 124

Qy        161 FQWNDDRCNMKNNFIC 176
               |||:||| |   :|
Db        125 --WNDERCNKKKLALC 138


Search completed: September  9, 2004, 22:22:59
Job time : 21 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: September 9, 2004, 22:18:15 ; Search time 12 Seconds
(without alignments)
1622.853 Million cell updates/sec

Title: US-09-887-855-2
Perfect score: 2000
Sequence: 1 MRPGTALQAVLLAVLLVGLR.....PDQMGRSKESGWVENEIYGY 374

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	657	32.9	273	1	CHOD_MOUSE	Q9cxm0 mus musculu
2	645	32.2	273	1	CHOD_HUMAN	Q9h9p2 homo sapien
3	183	9.2	1456	1	MANR_HUMAN	P22897 homo sapien
4	177	8.8	1268	1	PGCN_MOUSE	P55066 mus musculu
5	176	8.8	1321	1	PGCN_HUMAN	O14594 homo sapien
6	174.5	8.7	3381	1	PGCV_BOVIN	P81282 bos taurus
7	174	8.7	1257	1	PGCN_RAT	P55067 rattus norv
8	174	8.7	2738	1	PGCV_RAT	Q9erb4 rattus norv
9	174	8.7	3358	1	PGCV_MOUSE	Q62059 mus musculu
10	174	8.7	3396	1	PGCV_HUMAN	P13611 homo sapien
11	172.5	8.6	643	1	CD93_RAT	Q9et61 rattus norv
12	171	8.6	3562	1	PGCV_CHICK	Q90953 gallus gall
13	158.5	7.9	2364	1	PGCA_BOVIN	P13608 bos taurus
14	158.5	7.9	2415	1	PGCA_HUMAN	P16112 homo sapien
15	157.5	7.9	652	1	CD93_HUMAN	Q9npy3 homo sapien
16	155.5	7.8	2333	1	PGCA_CANFA	Q28343 canis famil
17	154.5	7.7	612	1	LEM2_MOUSE	Q00690 mus musculu
18	154.5	7.7	644	1	CD93_MOUSE	O89103 mus musculu
19	153.5	7.7	158	1	LECG_TRIST	Q9ygp1 trimeresuru
20	153.5	7.7	2124	1	PGCA_RAT	P07897 rattus norv
21	153	7.6	162	1	LEC3_MEGRO	P07439 megabalanus
22	152	7.6	912	1	PGCB_BOVIN	Q28062 bos taurus
23	151	7.5	197	1	CLE1_HUMAN	O75596 homo sapien
24	149	7.4	321	1	FCE2_HUMAN	P06734 homo sapien
25	148.5	7.4	152	1	IXA_TRIFL	P23806 trimeresuru
26	148.5	7.4	2132	1	PGCA_MOUSE	Q61282 mus musculu
27	147	7.3	883	1	PGCB_MOUSE	Q61361 mus musculu
28	146	7.3	2109	1	PGCA_CHICK	P07898 gallus gall
29	145	7.2	883	1	PGCB_RAT	P55068 rattus norv
30	144.5	7.2	173	1	LEC2_MEGRO	P17346 megabalanus
31	144.5	7.2	372	1	LEM1_RAT	P30836 rattus norv
32	143.5	7.2	372	1	LEM1_MOUSE	P18337 mus musculu
33	143	7.1	549	1	LEM2_RAT	P98105 rattus norv
34	141.5	7.1	165	1	LIT1_MOUSE	P43137 mus musculu
35	141.5	7.1	331	1	FCE2_MOUSE	P20693 mus musculu
36	140	7.0	370	1	LEM1_BOVIN	P98131 bos taurus
37	139	7.0	248	1	PSPA_HUMAN	P07714 homo sapien
38	139	7.0	283	1	LECA_SARPE	P05047 sarcophaga
39	138.5	6.9	172	1	LECA_PLEWA	Q02988 pleurodeles
40	138.5	6.9	202	1	TETN_MOUSE	P43025 mus musculu
41	137.5	6.9	372	1	LEM1_MACMU	Q95198 macaca mula
42	137.5	6.9	372	1	LEM1_PAPHA	Q28768 papio hamad
43	136	6.8	175	1	LITH_BOVIN	P23132 bos taurus
44	134.5	6.7	132	1	ACAL_ANSAN	P83300 anser anser
45	134.5	6.7	372	1	LEM1_PONPY	Q95235 pongo pygma

ALIGNMENTS

RESULT 1
CHOD_MOUSE
ID CHOD_MOUSE STANDARD; PRT; 273 AA.
AC Q9CXM0; Q8VI31;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Chondrolectin precursor (Transmembrane protein MT75).
GN CHODL.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J;
RA Weng L., Smits P., Hubner R., Wouters J., Merregaert J.;
RT "Mt75, a low expressed c-type lectin gene involving in
RT chondrogenesis.";
RL Submitted (OCT-2000) to the EMBL/GenBank/DDBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Embryonic head;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
CC -!- SUBCELLULAR LOCATION: Type I membrane protein (Potential).
CC -!- SIMILARITY: Contains 1 C-type lectin family domain.
CC ---

CC ---
DR EMBL; AF311699; AAL50354.1; -.

```
DR   EMBL; AK014255; BAB29226.1; -.
DR   HSSP; P22897; 1EGG.
DR   MGD; MGI:2179069; Chodl.
DR   InterPro; IPR001304; Lectin_C.
DR   Pfam; PF00059; lectin_c; 1.
DR   SMART; SM00034; CLECT; 1.
DR   PROSITE; PS00615; C_TYPE_LECTIN_1; FALSE_NEG.
DR   PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
KW   Lectin; Transmembrane; Glycoprotein; Signal.
FT   SIGNAL        1     21       POTENTIAL.
FT   CHAIN        22    273       CHONDROLECTIN.
FT   DOMAIN       22    216       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    217    237       POTENTIAL.
FT   DOMAIN      238    273       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN       35    179       C-TYPE LECTIN.
FT   CARBOHYD     86     86       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CONFLICT     24     24       V -> W (IN REF. 2).
FT   CONFLICT    179    179       T -> K (IN REF. 2).
SQ   SEQUENCE   273 AA;  30303 MW;  E052D933F244F4C7 CRC64;

  Query Match             32.9%;  Score 657;  DB 1;  Length 273;
  Best Local Similarity   49.1%;  Pred. No. 3.7e-48;
  Matches  139;  Conservative   38;  Mismatches   80;  Indels   26;  Gaps    9;

Qy         10 VLLAVLLVGLRAATGRLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQ 69
              :| | ||    |   |::||| ||    : ||||: |||: | |::|:||: ||  :||
Db          8 LLGAALLCAQGAFARRVVSGQKVCFADVKHPCYKMAYFHELSSRVSFQEARLACESEGGV 67

Qy         70 LVSIESEDEQKLIEKFIENLLP-----SDGDFWIGLRRREEKQSNSTACQDLYAWTDGSI 124
              |:|:|:| ||||||  ::||       |||||||| |  : |: | || ||| |:|||
Db         68 LLSLENEAEQKLIESMLQNLTKPGTGISDGDFWIGLLRSGDGQT-SGACPDLYQWSDGSS 126

Qy        125 SQFRNWYVDEPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDE-KP 183
              ||||||| |||||||| ||||||||:|  |:||||::|||||||||:|:|| |   |  |
Db        127 SQFRNWYTDEPSCGSEKCVVMYHQPTANPGLGGPYLYQWNDDRCNMKHNYICTYEPEIHP 186

Qy        184 AVPSREAEGEETELTTPVLPEETQEEDAKKTFKESREAAL--NLAYILIPSIPLLLLVV 241
                |:      |:  ||    |||| |            || :  || |::||:||||||::|
Db        187 TEPA-----EKPYLTNQ--PEETHENVV------VTEAGIIPNLIYVIIPTIPLLLLILV 233

Qy        242 TTVVCWVWICRKRK-REQPDPSTKKQHTIWPSPHQGNSPDLEV 283
                  |   :  | | | :  |:    | |:| |         :|
Db        234 ALGTCCFQMLHKSKGRSKTSPN---QSTLWISKSTRKESGMEV 273


RESULT 2
CHOD_HUMAN
ID   CHOD_HUMAN     STANDARD;      PRT;   273 AA.
AC   Q9H9P2; Q9HCY3;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   Chondrolectin precursor (Transmembrane protein MT75) (Protein PRED12).
GN   CHODL OR C21ORF68.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.
RX   MEDLINE=22074930; PubMed=12079284;
RA   Weng L., Smits P., Wauters J., Merregaert J.;
RT   "Molecular cloning and characterization of human chondrolectin, a
RT   novel type I transmembrane protein homologous to C-type lectins.";
RL   Genomics 80:62-70(2002).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20289799; PubMed=10830953;
RA   Hattori M., Fujiyama A., Taylor T.D., Watanabe H., Yada T.,
RA   Park H.-S., Toyoda A., Ishii K., Totoki Y., Choi D.-K., Groner Y.,
RA   Soeda E., Ohki M., Takagi T., Sakaki Y., Taudien S., Blechschmidt K.,
RA   Polley A., Menzel U., Delabar J., Kumpf K., Lehmann R., Patterson D.,
RA   Reichwald K., Rump A., Schillhabel M., Schudy A., Zimmermann W.,
RA   Rosenthal A., Kudoh J., Shibuya K., Kawasaki K., Asakawa S.,
RA   Shintani A., Sasaki T., Nagamine K., Mitsuyama S., Antonarakis S.E.,
RA   Minoshima S., Shimizu N., Nordsiek G., Hornischer K., Brandt P.,
RA   Scharfe M., Schoen O., Desario A., Reichelt J., Kauer G., Bloecker H.,
RA   Ramser J., Beck A., Klages S., Hennig S., Riesselmann L., Dagand E.,
RA   Wehrmeyer S., Borzym K., Gardiner K., Nizetic D., Francis F.,
RA   Lehrach H., Reinhardt R., Yaspo M.-L.;
RT   "The DNA sequence of human chromosome 21.";
RL   Nature 405:311-319(2000).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Brain;
RX   MEDLINE=22388257; PubMed=12477932;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length
RT   human and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [4]
RP   SEQUENCE OF 27-273 FROM N.A.
RA   Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
RA   Nishikawa T., Nagai K., Sugano S., Shiratori A., Sudo H.,
RA   Wagatsuma M., Hosoiri T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,
RA   Takahashi M., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,
RA   Watanabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K.,
RA   Yamamoto J., Wakamatsu A., Nakamura Y., Nagahari K., Masuho Y.,
RA   Ninomiya K., Iwayanagi T.;
RT   "NEDO human cDNA sequencing project.";
RL   Submitted (AUG-2000) to the EMBL/GenBank/DDBJ databases.
RN   [5]
RP   TISSUE SPECIFICITY.
RX   MEDLINE=21564202; PubMed=11707072;
RA   Reymond A., Friedli M., Neergard Henrichsen C., Chapot F., Deutsch S.,
RA   Ucla C., Rossier C., Lyle R., Guipponi M., Antonarakis S.E.;
RT   "From PREDs and open reading frames to cDNA isolation: revisiting the
RT   human chromosome 21 transcription map.";
RL   Genomics 78:46-54(2001).
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein (Potential).
CC   -!- TISSUE SPECIFICITY: Found in spleen, testis, prostate and fetal
CC       liver. Expression limited to vascular muscle of testis, smooth
CC       muscle of prostate stroma, heart muscle, skeletal muscle, crypts
CC       of small intestine, and red pulp of spleen.
CC   -!- PTM: N-glycosylated.
CC   -!- SIMILARITY: Contains 1 C-type lectin family domain.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; AF257472; AAL05981.1; -.
DR   EMBL; AL163217; CAB90388.1; -.
DR   EMBL; BC009418; AAH09418.1; -.
DR   EMBL; AK022689; BAB14181.1; ALT_INIT.
DR   HSSP; P22897; 1EGG.
```

```
DR   Genew; HGNC:17807; CHODL.
DR   MIM; 607247; -.
DR   InterPro; IPR001304; Lectin_C.
DR   Pfam; PF00059; lectin_c; 1.
DR   SMART; SM00034; CLECT; 1.
DR   PROSITE; PS00615; C_TYPE_LECTIN_1; FALSE_NEG.
DR   PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
KW   Lectin; Transmembrane; Glycoprotein; Signal.
FT   SIGNAL        1    21       POTENTIAL.
FT   CHAIN        22   273       CHONDROLECTIN.
FT   DOMAIN       22   216       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    217   237       POTENTIAL.
FT   DOMAIN      238   273.      CYTOPLASMIC (POTENTIAL).
FT   DOMAIN       35   179       C-TYPE LECTIN.
FT   CARBOHYD     86    86       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   273 AA;  30431 MW;  F4890AAFB572A311 CRC64;

  Query Match             32.2%;  Score 645;  DB 1;  Length 273;
  Best Local Similarity   48.4%;  Pred. No. 3.8e-47;
  Matches  137;  Conservative   40;  Mismatches   80;  Indels   26;  Gaps    9;

Qy         10 VLLAVLLVGLRAATGRLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQ 69
              :| | || |  |    |::||| ||    : ||||: |||: | |::|:||: ||  :||
Db          8 LLGAALLCGHGAFCRRVVSGQKVCFADFKHPCYKMAYFHELSSRVSFQEARLACESEGGV 67

Qy         70 LVSIESEDEQKLIEKFIENLLP-----SDGDFWIGLRRREEKQSNSTACQDLYAWTDGSI 124
              |:|:|:| ||||||  ::||       |||||||| |  : |: | || ||| |:|||
Db         68 LLSLENEAEQKLIESMLQNLTKPGTGISDGDFWIGLWRNGDGQT-SGACPDLYQWSDGSN 126

Qy        125 SQFRNWYVDEPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDE-KP 183
              ||:|||| |||||||| |||||||:|  |:||||::|||||||||:|:|||| |  |
Db        127 SQYRNWYTDEPSCGSEKCVVMYHQPTANPGLGGPYLYQWNDDRCNMKHNYICKYEPEINP 186

Qy        184 AVPSREAEGEETELTTPVLPEETQEEDAKKTFKESREAAL--NLAYILIPSIPLLLLLVV 241
                |      |:  ||    | :| :           || :  || |::||:||||||::|
Db        187 TAPV-----EKPYLTNQ--PGDTHQNVV------VTEAGIIPNLIYVVIPTIPLLLLILV 233

Qy        242 TTVVCWVWICRKRK-REQPDPSTKKQHTIWPSPHQGNSPDLEV 283
                |   : | | | :  |:   | |:| |        :||
Db        234 AFGTCCFQMLHKSKGRTKTSPN---QSTLWISKSTRKESGMEV 273


RESULT 3
MANR_HUMAN
ID   MANR_HUMAN     STANDARD;      PRT;  1456 AA.
AC   P22897;
DT   01-AUG-1991 (Rel. 19, Created)
DT   01-AUG-1991 (Rel. 19, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   Macrophage mannose receptor precursor (MMR) (CD206 antigen).
GN   MRC1.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC   TISSUE=Placenta;
RX   MEDLINE=90324192; PubMed=2373685;
RA   Taylor M.E., Conary J.T., Lennartz M.R., Stahl P.D., Drickamer K.;
RT   "Primary structure of the mannose receptor contains multiple motifs
RT   resembling carbohydrate-recognition domains.";
RL   J. Biol. Chem. 265:12156-12162(1990).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=93052405; PubMed=1294118;
RA   Kim S.J., Ruiz N., Bezouska K., Drickamer K.;
RT   "Organization of the gene encoding the human macrophage mannose
RT   receptor (MRC1).";
RL   Genomics 14:721-727(1992).
RN   [3]
RP   STUDIES ON THE BINDING OF INDIVIDUAL LECTIN DOMAINS.
RX   MEDLINE=92112893; PubMed=1730714;
RA   Taylor M.E., Bezouska K., Drickamer K.;
RT   "Contribution to ligand binding by multiple carbohydrate-recognition
RT   domains in the macrophage mannose receptor.";
RL   J. Biol. Chem. 267:1719-1726(1992).
RN   [4]
RP   X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS) OF 642-788.
RX   MEDLINE=20347275; PubMed=10779515;
RA   Feinberg H., Park-Snyder S., Kolatkar A.R., Heise C.T., Taylor M.E.,
RA   Weis W.I.;
RT   "Structure of a C-type carbohydrate recognition domain from the
RT   macrophage mannose receptor.";
RL   J. Biol. Chem. 275:21539-21548(2000).
CC   -!- FUNCTION: Mediates the endocytosis of glycoproteins by
CC       macrophages, in several recognition and uptake processes.
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC   -!- MISCELLANEOUS: CRDs 1-3 have at most very weak affinity for
CC       carbohydrate. CRD 4 shows the highest affinity binding and has
CC       multispecificity for a variety of monosaccharides. At least 3 CRDs
CC       (4, 5, and 7) are required for high affinity binding and
CC       endocytosis of multivalent glycoconjugates.
CC   -!- SIMILARITY: Contains 8 C-type lectin family domains.
CC   -!- SIMILARITY: Contains 1 ricin B-type lectin domain.
CC   -!- DATABASE: NAME=PROW; NOTE=PROW 2:85-89(2001);
CC       WWW="http://www.ncbi.nlm.nih.gov/prow/guide/1644341535_g.htm".
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; J05550; AAA59868.1; -.
DR   EMBL; M93221; AAA60389.1; -.
DR   EMBL; M93192; AAA60389.1; JOINED.
DR   EMBL; M93193; AAA60389.1; JOINED.
DR   EMBL; M93194; AAA60389.1; JOINED.
DR   EMBL; M93195; AAA60389.1; JOINED.
DR   EMBL; M93196; AAA60389.1; JOINED.
DR   EMBL; M93197; AAA60389.1; JOINED.
DR   EMBL; M93198; AAA60389.1; JOINED.
DR   EMBL; M93199; AAA60389.1; JOINED.
DR   EMBL; M93200; AAA60389.1; JOINED.
DR   EMBL; M93201; AAA60389.1; JOINED.
DR   EMBL; M93202; AAA60389.1; JOINED.
DR   EMBL; M93203; AAA60389.1; JOINED.
DR   EMBL; M93204; AAA60389.1; JOINED.
DR   EMBL; M93205; AAA60389.1; JOINED.
DR   EMBL; M93206; AAA60389.1; JOINED.
DR   EMBL; M93207; AAA60389.1; JOINED.
DR   EMBL; M93208; AAA60389.1; JOINED.
DR   EMBL; M93209; AAA60389.1; JOINED.
DR   EMBL; M93210; AAA60389.1; JOINED.
DR   EMBL; M93211; AAA60389.1; JOINED.
DR   EMBL; M93212; AAA60389.1; JOINED.
DR   EMBL; M93213; AAA60389.1; JOINED.
DR   EMBL; M93214; AAA60389.1; JOINED.
DR   EMBL; M93215; AAA60389.1; JOINED.
DR   EMBL; M93216; AAA60389.1; JOINED.
DR   EMBL; M93217; AAA60389.1; JOINED.
DR   EMBL; M93218; AAA60389.1; JOINED.
DR   EMBL; M93219; AAA60389.1; JOINED.
DR   EMBL; M93220; AAA60389.1; JOINED.
DR   PIR; A36563; A36563.
DR   PDB; 1EGG; 30-AUG-00.
DR   PDB; 1EGI; 30-AUG-00.
DR   Genew; HGNC:7228; MRC1.
DR   MIM; 153618; -.
DR   GO; GO:0005887; C:integral to plasma membrane; TAS.
DR   GO; GO:0005537; F:mannose binding; TAS.
DR   GO; GO:0004872; F:receptor activity; TAS.
DR   GO; GO:0006898; P:receptor mediated endocytosis; TAS.
```

```
DR   InterPro; IPR002353; AntifreezeII.
DR   InterPro; IPR000562; FN_Type_II.
DR   InterPro; IPR001304; Lectin_C.
DR   InterPro; IPR000772; Ricin_B_lectin.
DR   InterPro; IPR008997; RicinB_like.
DR   Pfam; PF00040; fn2; 1.
DR   Pfam; PF00059; lectin_c; 8.
DR   Pfam; PF00652; Ricin_B_lectin; 2.
DR   PRINTS; PR00356; ANTIFREEZEII.
DR   PRINTS; PR00013; FNTYPEII.
DR   ProDom; PD000995; FN_Type_II; 1.
DR   SMART; SM00034; CLECT; 8.
DR   SMART; SM00059; FN2; 1.
DR   SMART; SM00458; RICIN; 1.
DR   PROSITE; PS00615; C_TYPE_LECTIN_1; 6.
DR   PROSITE; PS50041; C_TYPE_LECTIN_2; 8.
DR   PROSITE; PS00023; FIBRONECTIN_2; 1.
DR   PROSITE; PS50231; RICIN_B_LECTIN; 1.
KW   Receptor; Signal; Calcium-binding; Transmembrane; Repeat;
KW   Glycoprotein; Endocytosis; Macrophage; Lectin; Antigen; 3D-structure.
FT   SIGNAL        1     18       POTENTIAL.
FT   CHAIN        19   1456       MACROPHAGE MANNOSE RECEPTOR.
FT   DOMAIN       19   1383       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   1384   1411       POTENTIAL.
FT   DOMAIN     1412   1456       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN       22    142       RICIN B-TYPE LECTIN.
FT   DOMAIN      157    212       FIBRONECTIN TYPE-II.
FT   DOMAIN      216    344       C-TYPE LECTIN 1 (LONG FORM).
FT   DOMAIN      360    490       C-TYPE LECTIN 2 (LONG FORM).
FT   DOMAIN      502    629       C-TYPE LECTIN 3 (LONG FORM).
FT   DOMAIN      644    781       C-TYPE LECTIN 4 (LONG FORM).
FT   DOMAIN      805    926       C-TYPE LECTIN 5 (LONG FORM).
FT   DOMAIN      943   1083       C-TYPE LECTIN 6 (LONG FORM).
FT   DOMAIN     1100   1216       C-TYPE LECTIN 7 (LONG FORM).
FT   DOMAIN     1228   1359       C-TYPE LECTIN 8 (LONG FORM).
FT   DISULFID    646    659
FT   DISULFID    680    777
FT   DISULFID    753    769
FT   CARBOHYD    104    104       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    344    344       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    529    529       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    926    926       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    930    930       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1160   1160       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1205   1205       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1311   1311       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   TURN        648    649
FT   STRAND      651    652
FT   TURN        654    655
FT   STRAND      658    663
FT   HELIX       667    669
FT   STRAND      671    671
FT   HELIX       673    683
FT   TURN        684    684
FT   STRAND      686    687
FT   HELIX       693    705
FT   TURN        706    707
FT   TURN        709    710
FT   STRAND      712    718
FT   TURN        723    724
FT   STRAND      727    727
FT   TURN        729    730
FT   STRAND      733    733
FT   TURN        741    742
FT   HELIX       746    748
FT   STRAND      752    757
FT   TURN        758    761
FT   STRAND      764    768
FT   TURN        769    770
FT   STRAND      773    780
FT   TURN        781    782
SQ   SEQUENCE   1456 AA;  166011 MW;  264E5AF3C576A5E3 CRC64;
```

```
  Query Match             9.2%;  Score 183;  DB 1;  Length 1456;
  Best Local Similarity   25.1%;  Pred. No. 4.2e-07;
  Matches   53;  Conservative   38;  Mismatches   66;  Indels   54;  Gaps    8;

Qy         42 YKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFWIGLR 101
              ||   :: :  :   : |:  |:|: | ||||:|| |:| : |:: |   :   ::|||
Db        807 YKDYQYYFSKEKETMDNARAFCKRNFGDLVSIQSESEKKFLWKYV-NRNDAQSAYFIGLL 865

Qy        102 RREEKQSNSTACQDLYAWTDGSISQFRNWYVDEPSCGS--EVCVVMYHQPSAPAGIGGPY 159
                :|:          :|| |||   : :|   ||:  :  | || ||    : :|
Db        866 ISLDKK---------FAWMDGSKVDYVSWATGEPNFANEDENCVTMY----SNSGF---- 908

Qy        160 MFQWNDDRCNMKNNFICKYSDEK----PAVPSREAEGEETELTTPVLPEETQE------- 208
                |||  |   | |||:  :         :|             | | :|   :|
Db        909 ---WNDINCGYPNAFICQRHNSSINATTVMP----------TMPSVPSGCKEGWNFYSN 954

Qy        209 ---------EDAKKTFKESREAALNLAYILI 230
                      |: :| ::|:|:| :     |:
Db        955 KCFKIFGFMEEERKNWQEARKACIGFGGNLV 985


RESULT 4
PGCN_MOUSE
ID   PGCN_MOUSE     STANDARD;      PRT;  1268 AA.
AC   P55066;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Neurocan core protein precursor (Chondroitin sulfate proteoglycan 3).
GN   CSPG3 OR NCAN.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BALB/c; TISSUE=Brain;
RX   MEDLINE=96039250; PubMed=7490074;
RA   Rauch U., Grimpe B., Kulbe G., Arnold-Ammer I., Beier D.,
RA   Faessler R.;
RT   "Structure and chromosomal localization of the mouse neurocan gene.";
RL   Genomics 28:405-410(1995).
CC   -!- FUNCTION: May modulate neuronal adhesion and neurite growth during
CC       development by binding to neural cell adhesion molecules (NG-CAM
CC       and N-CAM). Chondroitin sulfate proteoglycan; binds to hyaluronic
CC       acid.
CC   -!- TISSUE SPECIFICITY: Brain.
CC   -!- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
CC   -!- SIMILARITY: Contains 2 EGF-like domains.
CC   -!- SIMILARITY: Contains 2 link domains.
CC   -!- SIMILARITY: Contains 1 C-type lectin family domain.
CC   -!- SIMILARITY: Contains 1 Sushi (SCR) domain.
CC   -!- SIMILARITY: Belongs to the aggrecan/versican proteoglycan family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X84727; CAA59216.1; -.
DR   PIR; S52781; S52781.
DR   HSSP; P00740; 1EDM.
DR   MGD; MGI:104694; Cspg3.
DR   InterPro; IPR002353; AntifreezeII.
DR   InterPro; IPR000152; Asx_hydroxyl_S.
DR   InterPro; IPR000742; EGF_2.
DR   InterPro; IPR001881; EGF_Ca.
DR   InterPro; IPR006209; EGF_like.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003599; Ig.
```

```
DR   InterPro; IPR001304; Lectin_C.
DR   InterPro; IPR000538; Link.
DR   InterPro; IPR000436; Sushi_SCR_CCP.
DR   Pfam; PF00008; EGF; 2.
DR   Pfam; PF00047; ig; 1.
DR   Pfam; PF00059; lectin_c; 1.
DR   Pfam; PF00084; sushi; 1.
DR   Pfam; PF00193; Xlink; 2.
DR   PRINTS; PR00356; ANTIFREEZEII.
DR   PRINTS; PR01265; LINKMODULE.
DR   ProDom; PD000918; Link; 2.
DR   SMART; SM00032; CCP; 1.
DR   SMART; SM00034; CLECT; 1.
DR   SMART; SM00179; EGF_CA; 1.
DR   SMART; SM00409; IG; 1.
DR   SMART; SM00445; LINK; 2.
DR   PROSITE; PS00010; ASX_HYDROXYL; 1.
DR   PROSITE; PS00615; C_TYPE_LECTIN_1; 1.
DR   PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
DR   PROSITE; PS00022; EGF_1; 3.
DR   PROSITE; PS01186; EGF_2; 1.
DR   PROSITE; PS50026; EGF_3; 2.
DR   PROSITE; PS01187; EGF_CA; 1.
DR   PROSITE; PS50835; IG_LIKE; 1.
DR   PROSITE; PS01241; LINK; 2.
KW   Glycoprotein; Hyaluronic acid; Proteoglycan; Immunoglobulin domain;
KW   EGF-like domain; Calcium; Repeat; Lectin; Sushi; Signal.
FT   SIGNAL        1     22       POTENTIAL.
FT   CHAIN        23   1268       NEUROCAN CORE PROTEIN.
FT   DOMAIN       37    157       IG-LIKE V-TYPE.
FT   DOMAIN      158    253       LINK 1.
FT   DOMAIN      259    355       LINK 2.
FT   DOMAIN      960    996       EGF-LIKE 1.
FT   DOMAIN      998   1034       EGF-LIKE 2, CALCIUM-BINDING (POTENTIAL).
FT   DOMAIN     1036   1165       C-TYPE LECTIN.
FT   DOMAIN     1166   1224       SUSHI.
FT   DISULFID     58    139       BY SIMILARITY.
FT   DISULFID    181    252       BY SIMILARITY.
FT   DISULFID    205    226       BY SIMILARITY.
FT   DISULFID    279    354       BY SIMILARITY.
FT   DISULFID    303    324       BY SIMILARITY.
FT   DISULFID    964    975       BY SIMILARITY.
FT   DISULFID    969    984       BY SIMILARITY.
FT   DISULFID    986    995       BY SIMILARITY.
FT   DISULFID   1040   1051       BY SIMILARITY.
FT   DISULFID   1068   1160       BY SIMILARITY.
FT   DISULFID   1136   1152       BY SIMILARITY.
FT   DISULFID   1167   1210       BY SIMILARITY.
FT   DISULFID   1196   1223       BY SIMILARITY.
FT   CARBOHYD    121    121       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    339    339       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    742    742       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    978    978       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1175   1175       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   1268 AA;  137200 MW;  3014E8E202A2FAEC CRC64;

  Query Match             8.8%;  Score 177;  DB 1;  Length 1268;
  Best Local Similarity   31.5%;  Pred. No. 1.1e-06;
  Matches   45;  Conservative   17;  Mismatches   49;  Indels   32;  Gaps    6;

Qy         38 QRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFW 97
              |  ||:  ||   : |  :|:|:  |||  | | |: | :| | |  |       : |
Db       1048 QGHCYR--YF---AHRRAWEDAERDCRRRAGHLTSVHSPEEHKFINSF------GHENSW 1096

Qy         98 IGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEPS---CGSEVCVVMYHQPSAPAG 154
              |||  |  ::         : ||| :  |: ||   :|     | | ||||    |
Db       1097 IGLNDRTVERD--------FQWTDNTGLQYENWREKQPDNFFAGGEDCVVMVAHESG--- 1145

Qy        155 IGGPYMFQWNDDRCNMKNNFICK 177
                    :|||  ||    ::||
Db       1146 -------RWNDVPCNYNLPYVCK 1161
```

```
RESULT 5
PGCN_HUMAN
ID   PGCN_HUMAN     STANDARD;      PRT;  1321 AA.
AC   O14594; Q9UPK6;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Neurocan core protein precursor (Chondroitin sulfate proteoglycan 3).
GN   CSPG3 OR NCAN OR NEUR.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=99013874; PubMed=9795216;
RA   Prange C.K., Pennacchio L.A., Lieuallen K., Fan W., Lennon G.G.;
RT   "Characterization of the human neurocan gene, CSPG3.";
RL   Gene 221:199-205(1998).
RN   [2]
RP   SEQUENCE OF 1-990 AND 1007-1321 FROM N.A., AND VARIANT ALA-1254.
RA   Lamerdin J.E., McCready P.M., Skowronski E., Adamson A.W.,
RA   Burkhart-Schultz K., Gordon L., Kyle A., Ramirez M., Stilwagen S.,
RA   Phan H., Velasco N., Do L., Regala W., Terry A., Garnes J.,
RA   Danganan L., Poundstone P., Christensen M., Georgescu A., Avila J.,
RA   Liu S., Attix C., Andreise T., Trankheim M., Amico-Keller G.,
RA   Coefield J., Duarte S., Lucas S., Bruce R., Thomas P., Quan G.,
RA   Kronmiller B., Arellano A., Montgomery M., Ow D., Nolan M., Trong S.,
RA   Kobayashi A., Olsen A.S., Carrano A.V.;
RT   "Sequence analysis of an ~1 Mb region containing the MEF2B gene in
RT   19p12.";
RL   Submitted (JUL-1998) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: May modulate neuronal adhesion and neurite growth during
CC       development by binding to neural cell adhesion molecules (NG-CAM
CC       and N-CAM). Chondroitin sulfate proteoglycan; binds to hyaluronic
CC       acid.
CC   -!- TISSUE SPECIFICITY: Brain.
CC   -!- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
CC   -!- SIMILARITY: Contains 2 EGF-like domains.
CC   -!- SIMILARITY: Contains 2 link domains.
CC   -!- SIMILARITY: Contains 1 C-type lectin family domain.
CC   -!- SIMILARITY: Contains 1 Sushi (SCR) domain.
CC   -!- SIMILARITY: Belongs to the aggrecan/versican proteoglycan family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF026547; AAC80576.1; -.
DR   EMBL; AC003110; AAB86655.1; -.
DR   EMBL; AC005254; AAC25581.1; -.
DR   HSSP; P00740; 1EDM.
DR   Genew; HGNC:2465; CSPG3.
DR   MIM; 600826; -.
DR   InterPro; IPR000152; Asx_hydroxyl_S.
DR   InterPro; IPR000742; EGF_2.
DR   InterPro; IPR001881; EGF_Ca.
DR   InterPro; IPR006209; EGF_like.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003599; Ig.
DR   InterPro; IPR001304; Lectin_C.
DR   InterPro; IPR000538; Link.
DR   InterPro; IPR000436; Sushi_SCR_CCP.
DR   Pfam; PF00008; EGF; 2.
DR   Pfam; PF00047; ig; 1.
DR   Pfam; PF00059; lectin_c; 1.
DR   Pfam; PF00084; sushi; 1.
DR   Pfam; PF00193; Xlink; 2.
DR   PRINTS; PR01265; LINKMODULE.
```

```
DR   ProDom; PD000918; Link; 2.
DR   SMART; SM00032; CCP; 1.
DR   SMART; SM00034; CLECT; 1.
DR   SMART; SM00179; EGF_CA; 1.
DR   SMART; SM00409; IG; 1.
DR   SMART; SM00445; LINK; 2.
DR   PROSITE; PS00010; ASX_HYDROXYL; 1.
DR   PROSITE; PS00615; C_TYPE_LECTIN_1; 1.
DR   PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
DR   PROSITE; PS00022; EGF_1; 3.
DR   PROSITE; PS01186; EGF_2; 1.
DR   PROSITE; PS50026; EGF_3; 2.
DR   PROSITE; PS01187; EGF_CA; 1.
DR   PROSITE; PS50835; IG_LIKE; 1.
DR   PROSITE; PS01241; LINK; 2.
KW   Glycoprotein; Hyaluronic acid; Proteoglycan; Immunoglobulin domain;
KW   EGF-like domain; Calcium; Repeat; Lectin; Sushi; Signal; Polymorphism.
FT   SIGNAL       1     22       POTENTIAL.
FT   CHAIN       23   1321       NEUROCAN CORE PROTEIN.
FT   DOMAIN      38    153       IG-LIKE V-TYPE.
FT   DOMAIN     159    254       LINK 1.
FT   DOMAIN     260    356       LINK 2.
FT   DOMAIN    1008   1044       EGF-LIKE 1.
FT   DOMAIN    1046   1082       EGF-LIKE 2, CALCIUM-BINDING (POTENTIAL).
FT   DOMAIN    1084   1213       C-TYPE LECTIN.
FT   DOMAIN    1214   1272       SUSHI.
FT   DISULFID    59    140       BY SIMILARITY.
FT   DISULFID   182    253       BY SIMILARITY.
FT   DISULFID   206    227       BY SIMILARITY.
FT   DISULFID   280    355       BY SIMILARITY.
FT   DISULFID   304    325       BY SIMILARITY.
FT   DISULFID  1012   1023       BY SIMILARITY.
FT   DISULFID  1017   1032       BY SIMILARITY.
FT   DISULFID  1034   1043       BY SIMILARITY.
FT   DISULFID  1088   1099       BY SIMILARITY.
FT   DISULFID  1116   1208       BY SIMILARITY.
FT   DISULFID  1184   1200       BY SIMILARITY.
FT   DISULFID  1215   1258       BY SIMILARITY.
FT   DISULFID  1244   1271       BY SIMILARITY.
FT   CARBOHYD   122    122       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   340    340       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD  1026   1026       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD  1223   1223       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   VARIANT   1254   1254       V -> A (in dbSNP:1064389).
FT                               /FTId=VAR_016176.
FT   CONFLICT  1234   1234       N -> Y (IN REF. 2).
FT   CONFLICT  1282   1282       G -> R (IN REF. 2).
SQ   SEQUENCE  1321 AA;  142972 MW;  2EF47F823DB980B8 CRC64;

  Query Match             8.8%;  Score 176;  DB 1;  Length 1321;
  Best Local Similarity   31.5%;  Pred. No. 1.4e-06;
  Matches   45;  Conservative   18;  Mismatches   48;  Indels   32;  Gaps    6;

Qy         38 QRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFW 97
              |  ||:  ||   : |  :|:|:: |||  | | |: | :|   |  |         : |
Db       1096 QGHCYR--YF---AHRRAWEDAEKDCRRRSGHLTSVHSPEEHSFINSF------GHENTW 1144

Qy         98 IGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEPS---CGSEVCVVMYHQPSAPAG 154
              |||  |  ::         : ||| :  || ||  ::|     | | ||||    |
Db       1145 IGLNDRIVERD--------FQWTDNTGLQFENWRENQPDNFFAGGEDCVVMVAHESG--- 1193

Qy        155 IGGPYMFQWNDDRCNMKNNFICK 177
                 :|||  ||    ::||
Db       1194 -------RWNDVPCNYNLPYVCK 1209


RESULT 6
PGCV_BOVIN
ID   PGCV_BOVIN     STANDARD;      PRT;  3381 AA.
AC   P81282; O77609; O77610; O77611; O77612;
DT   15-DEC-1998 (Rel. 37, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Versican core protein precursor (Large fibroblast proteoglycan)
DE   (Chondroitin sulfate proteoglycan core protein 2) (PG-M) (Glial
DE   hyaluronate-binding protein) (GHAP).
GN   CPSG2.
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORMS V0; V1; V2 AND V3).
RC   TISSUE=Forebrain;
RX   MEDLINE=98288320; PubMed=9624174;
RA   Schmalfeldt M., Dours-Zimmermann M.T., Winterhalter K.H.,
RA   Zimmermann D.R.;
RT   "Versican V2 is a major extracellular matrix component of the mature
RT   bovine brain.";
RL   J. Biol. Chem. 273:15758-15764(1998).
RN   [2]
RP   SEQUENCE OF 21-53; 78-96; 226-250; 262-277; 295-306; 314-324; 329-331
RP   AND 342-348.
RC   TISSUE=Spinal cord;
RX   MEDLINE=92062692; PubMed=1720020;
RA   Perides G., Biviano F., Bignami A.;
RT   "Interaction of a brain extracellular matrix protein with hyaluronic
RT   acid.";
RL   Biochim. Biophys. Acta 1075:248-258(1991).
CC   -!- FUNCTION: May play a role in intercellular signaling and in
CC       connecting cells with the extracellular matrix. May take part in
CC       the regulation of cell motility, growth and differentiation. Binds
CC       hyaluronic acid.
CC   -!- SUBUNIT: Interacts with FBLN1 (By similarity).
CC   -!- SUBCELLULAR LOCATION: Secreted; extracellular matrix.
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=4;
CC         Comment=Additional isoforms seem to exist;
CC       Name=V0;
CC         IsoId=P81282-1; Sequence=Displayed;
CC       Name=V1;
CC         IsoId=P81282-2; Sequence=VSP_003078, VSP_003079;
CC       Name=V2;
CC         IsoId=P81282-3; Sequence=VSP_003080;
CC       Name=V3;
CC         IsoId=P81282-4; Sequence=VSP_003078, VSP_003081;
CC   -!- TISSUE SPECIFICITY: Cerebral white matter. V0 and V1 are expressed
CC       in the central nervous system, and in a number of mesenchymal and
CC       epithelial tissues; the major isoform v2 is restricted to the
CC       central nervous system.
CC   -!- DEVELOPMENTAL STAGE: Disappears after the cartilage development
CC       (By similarity).
CC   -!- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
CC   -!- SIMILARITY: Contains 2 link domains.
CC   -!- SIMILARITY: Contains 2 EGF-like domains.
CC   -!- SIMILARITY: Contains 1 C-type lectin family domain.
CC   -!- SIMILARITY: Contains 1 Sushi (SCR) domain.
CC   -!- SIMILARITY: Belongs to the aggrecan/versican proteoglycan family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF060456; AAC24358.1; -.
DR   EMBL; AF060457; AAC24359.1; -.
DR   EMBL; AF060458; AAC24360.1; -.
DR   EMBL; AF060459; AAC24361.1; -.
DR   PIR; T14274; T14274.
DR   PIR; T42389; T42389.
DR   HSSP; P01132; 1EPG.
DR   InterPro; IPR000152; Asx_hydroxyl_S.
```

```
DR   InterPro; IPR000742; EGF_2.
DR   InterPro; IPR001881; EGF_Ca.
DR   InterPro; IPR006209; EGF_like.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003599; Ig.
DR   InterPro; IPR001304; Lectin_C.
DR   InterPro; IPR000538; Link.
DR   InterPro; IPR000436; Sushi_SCR_CCP.
DR   Pfam; PF00008; EGF; 2.
DR   Pfam; PF00047; ig; 1.
DR   Pfam; PF00059; lectin_c; 1.
DR   Pfam; PF00084; sushi; 1.
DR   Pfam; PF00193; Xlink; 2.
DR   PRINTS; PR01265; LINKMODULE.
DR   ProDom; PD000918; Link; 2.
DR   SMART; SM00032; CCP; 1.
DR   SMART; SM00034; CLECT; 1.
DR   SMART; SM00179; EGF_CA; 1.
DR   SMART; SM00409; IG; 1.
DR   SMART; SM00445; LINK; 2.
DR   PROSITE; PS00010; ASX_HYDROXYL; 1.
DR   PROSITE; PS00615; C_TYPE_LECTIN_1; 1.
DR   PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
DR   PROSITE; PS00022; EGF_1; 2.
DR   PROSITE; PS01186; EGF_2; 1.
DR   PROSITE; PS50026; EGF_3; 2.
DR   PROSITE; PS01187; EGF_CA; 1.
DR   PROSITE; PS50835; IG_LIKE; 1.
DR   PROSITE; PS01241; LINK; 2.
KW   Glycoprotein; Proteoglycan; Lectin; Extracellular matrix; Sushi;
KW   Signal; Repeat; EGF-like domain; Calcium; Immunoglobulin domain;
KW   Hyaluronic acid; Alternative splicing.
FT   SIGNAL        1     20       POTENTIAL.
FT   CHAIN        21   3381       VERSICAN CORE PROTEIN.
FT   DOMAIN       21    147       IG-LIKE V-TYPE.
FT   DOMAIN      168    245       LINK 1.
FT   DOMAIN      266    347       LINK 2.
FT   DOMAIN      349   1336       GAG-ALPHA
FT                                (GLUCOSAMINOGLYCAN ATTACHMENT DOMAIN).
FT   DOMAIN     1337   3074       GAG-BETA.
FT   DOMAIN     3074   3110       EGF-LIKE 1.
FT   DOMAIN     3112   3148       EGF-LIKE 2, CALCIUM-BINDING.
FT   DOMAIN     3161   3275       C-TYPE LECTIN.
FT   DOMAIN     3280   3338       SUSHI.
FT   DISULFID     44    131       BY SIMILARITY.
FT   DISULFID    173    244       BY SIMILARITY.
FT   DISULFID    197    218       BY SIMILARITY.
FT   DISULFID    271    346       BY SIMILARITY.
FT   DISULFID    295    316       BY SIMILARITY.
FT   DISULFID   3078   3089       BY SIMILARITY.
FT   DISULFID   3083   3098       BY SIMILARITY.
FT   DISULFID   3100   3109       BY SIMILARITY.
FT   DISULFID   3116   3127       BY SIMILARITY.
FT   DISULFID   3121   3136       BY SIMILARITY.
FT   DISULFID   3138   3147       BY SIMILARITY.
FT   DISULFID   3154   3165       BY SIMILARITY.
FT   DISULFID   3182   3274       BY SIMILARITY.
FT   DISULFID   3250   3266       BY SIMILARITY.
FT   DISULFID   3281   3324       BY SIMILARITY.
FT   DISULFID   3310   3337       BY SIMILARITY.
FT   CARBOHYD     57     57       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    331    331       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    352    352       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    817    817       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    965    965       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1017   1017       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1333   1333       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1393   1393       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1437   1437       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1463   1463       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1653   1653       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1974   1974       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   2045   2045       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   2074   2074       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   2103   2103       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   2263   2263       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   2290   2290       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   2356   2356       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   2623   2623       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   2641   2641       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   2919   2919       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   3052   3052       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   3354   3354       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   3364   3364       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   VARSPLIC    349    349       P -> R (in isoform V1 and isoform V3).
FT                                /FTId=VSP_003078.
FT   VARSPLIC    350   1336       Missing (In isoform V1).
FT                                /FTId=VSP_003079.
FT   VARSPLIC   1337   3074       Missing (In isoform V2).
FT                                /FTId=VSP_003080.
FT   VARSPLIC    350   3074       Missing (In isoform V3).
FT                                /FTId=VSP_003081.
FT   CONFLICT     25     25       MISSING (IN REF. 2).
FT   CONFLICT     51     51       MISSING (IN REF. 2).
FT   CONFLICT     89     89       N -> D (IN REF. 2).
FT   CONFLICT     96     96       Q -> D (IN REF. 2).
FT   CONFLICT    346    346       C -> R (IN REF. 2).
SQ   SEQUENCE   3381 AA;  369984 MW;  F09716FA7778D459 CRC64;

  Query Match             8.7%;  Score 174.5;  DB 1;  Length 3381;
  Best Local Similarity   25.5%;  Pred. No. 7e-06;
  Matches   59;  Conservative   26;  Mismatches   71;  Indels   75;  Gaps   11;

Qy         38 QRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDF- 96
              |  |||  ||   : |  :: |:  ||  |  | || | :||  : :         |:
Db       3162 QGQCYK--YF---AHRRTWDAAERECRLQGAHLTSILSHEEQMFVNRV-------GHDYQ 3209

Qy         97 WIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEP----SCGSEVCVVMYHQPSAP 152
              ||||          |    :  : |||||  |: ||  ::|    | | :  |:::|:
Db       3210 WIGL--------NDKMFEHDFRWTDGSTLQYENWRPNQPDSFFSTGEDCVVIIWHENG-- 3259

Qy        153 AGIGGPYMFQWNDDRCNMKNNFICKYS----DEKPAVPSREAEGE--------------- 193
                 ||||  ||    : ||       : | | : :  |:
Db       3260 ---------QWNDVPCNYHLTYTCKKGTVACGQPPVVENAKTFGKMKPRYEINSLIRYHC 3310

Qy        194 -----ETELTT-----------PVL----PEETQEEDAKKTFKESREAALN 224
                    :  | |           | :    |   |   :|| || |  |  |
Db       3311 KDGFIQRHLPTIRCLGNGRWAMPKITCLNPSAYQRTYSKKYFKNSSSAKDN 3361


RESULT 7
PGCN_RAT
ID   PGCN_RAT       STANDARD;      PRT;  1257 AA.
AC   P55067;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   Neurocan core protein precursor (Chondroitin sulfate proteoglycan 3)
DE   (245 kDa early postnatal core glycoprotein) [Contains: 150 kDa adult
DE   core glycoprotein].
GN   CSPG3 OR NCAN.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC   STRAIN=Sprague-Dawley; TISSUE=Brain;
RX   MEDLINE=92406907; PubMed=1326557;
RA   Rauch U., Karthikeyan L., Maurel P., Margolis R.U., Margolis R.K.;
RT   "Cloning and primary structure of neurocan, a developmentally
RT   regulated, aggregating chondroitin sulfate proteoglycan of brain.";
RL   J. Biol. Chem. 267:19536-19547(1992).
RN   [2]
RP   CHARACTERIZATION.
RX   MEDLINE=94230574; PubMed=7513709;
```

```
RA   Friedlander D.R., Milev P., Karthikeyan L., Margolis R.K.,
RA   Margolis R.U., Grumet M.;
RT   "The neuronal chondroitin sulfate proteoglycan neurocan binds to the
RT   neural cell adhesion molecules Ng-CAM/L1/NILE and N-CAM, and inhibits
RT   neuronal adhesion and neurite outgrowth.";
RL   J. Cell Biol. 125:669-680(1994).
CC   -!- FUNCTION: May modulate neuronal adhesion and neurite growth during
CC       development by binding to neural cell adhesion molecules (NG-CAM
CC       and N-CAM). Chondroitin sulfate proteoglycan; binds to hyaluronic
CC       acid.
CC   -!- TISSUE SPECIFICITY: Early postnatal and adult brain; not expressed
CC       in kidney, lung, liver and muscle.
CC   -!- PTM: CONTAINS MOSTLY CHONDROITIN SULFATE, BUT ALSO N-LINKED AND
CC       O-LINKED OLIGOSACCHARIDES (BY SIMILARITY).
CC   -!- PTM: TWO ISOFORMS WERE FOUND THAT PROBABLY ARISE BY PROTEOLYTIC
CC       DEGRADATION. THE LARGE ISOFORM IS PREDOMINANT IN EARLY POSTNATAL
CC       BRAIN, THE SMALL ISOFORM IS FOUND IN ADULT BRAIN.
CC   -!- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
CC   -!- SIMILARITY: Contains 2 EGF-like domains.
CC   -!- SIMILARITY: Contains 2 link domains.
CC   -!- SIMILARITY: Contains 1 C-type lectin family domain.
CC   -!- SIMILARITY: Contains 1 Sushi (SCR) domain.
CC   -!- SIMILARITY: Belongs to the aggrecan/versican proteoglycan family.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; M97161; AAC37679.1; -.
DR   PIR; S28764; S28764.
DR   HSSP; P00740; 1EDM.
DR   InterPro; IPR002353; AntifreezeII.
DR   InterPro; IPR000152; Asx_hydroxyl_S.
DR   InterPro; IPR000742; EGF_2.
DR   InterPro; IPR001881; EGF_Ca.
DR   InterPro; IPR006209; EGF_like.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003599; Ig.
DR   InterPro; IPR001304; Lectin_C.
DR   InterPro; IPR000538; Link.
DR   InterPro; IPR000436; Sushi_SCR_CCP.
DR   Pfam; PF00008; EGF; 2.
DR   Pfam; PF00047; ig; 1.
DR   Pfam; PF00059; lectin_c; 1.
DR   Pfam; PF00084; sushi; 1.
DR   Pfam; PF00193; Xlink; 2.
DR   PRINTS; PR00356; ANTIFREEZEII.
DR   PRINTS; PR01265; LINKMODULE.
DR   ProDom; PD000918; Link; 2.
DR   SMART; SM00032; CCP; 1.
DR   SMART; SM00034; CLECT; 1.
DR   SMART; SM00179; EGF_CA; 1.
DR   SMART; SM00409; IG; 1.
DR   SMART; SM00445; LINK; 2.
DR   PROSITE; PS00010; ASX_HYDROXYL; 1.
DR   PROSITE; PS00615; C_TYPE_LECTIN_1; 1.
DR   PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
DR   PROSITE; PS00022; EGF_1; 3.
DR   PROSITE; PS01186; EGF_2; 1.
DR   PROSITE; PS50026; EGF_3; 2.
DR   PROSITE; PS01187; EGF_CA; 1.
DR   PROSITE; PS50835; IG_LIKE; 1.
DR   PROSITE; PS01241; LINK; 2.
KW   Glycoprotein; Hyaluronic acid; Proteoglycan; Immunoglobulin domain;
KW   EGF-like domain; Calcium; Repeat; Lectin; Sushi; Signal.
FT   SIGNAL       1     22
FT   CHAIN       23   1257       NEUROCAN CORE PROTEIN.
FT   CHAIN      639   1257       150 kDa ADULT CORE GLYCOPROTEIN.
FT   DOMAIN      37    157       IG-LIKE V-TYPE.
FT   DOMAIN     158    253       LINK 1.
FT   DOMAIN     259    355       LINK 2.
FT   DOMAIN     949    985       EGF-LIKE 1.
FT   DOMAIN     987   1023       EGF-LIKE 2, CALCIUM-BINDING (POTENTIAL).
FT   DOMAIN    1025   1154       C-TYPE LECTIN.
FT   DOMAIN    1155   1213       SUSHI.
FT   DISULFID    58    139       BY SIMILARITY.
FT   DISULFID   181    252       BY SIMILARITY.
FT   DISULFID   205    226       BY SIMILARITY.
FT   DISULFID   279    354       BY SIMILARITY.
FT   DISULFID   303    324       BY SIMILARITY.
FT   DISULFID   953    964       BY SIMILARITY.
FT   DISULFID   958    973       BY SIMILARITY.
FT   DISULFID   975    984       BY SIMILARITY.
FT   DISULFID  1029   1040       BY SIMILARITY.
FT   DISULFID  1057   1149       BY SIMILARITY.
FT   DISULFID  1125   1141       BY SIMILARITY.
FT   DISULFID  1156   1199       BY SIMILARITY.
FT   DISULFID  1185   1212       BY SIMILARITY.
FT   CARBOHYD   121    121       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   339    339       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   737    737       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   944    944       O-LINKED (XYL. . .) (CHONDROITIN SULFATE).
FT   CARBOHYD   967    967       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD  1164   1164       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE  1257 AA;  135544 MW;  992B33DCFA19EE1B CRC64;

  Query Match             8.7%;  Score 174;  DB 1;  Length 1257;
  Best Local Similarity   30.8%;  Pred. No. 2e-06;
  Matches   44;  Conservative   18;  Mismatches   49;  Indels   32;  Gaps    6;

Qy         38 QRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFW 97
              |  ||:  ||   : |  :|:|:  |||  | | |: | :| | |  |          : |
Db       1037 QGHCYR--YF---AHRRAWEDAERDCRRRAGHLTSVHSPEEHKFINSF------GHENSW 1085

Qy         98 IGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEPS---CGSEVCVVMYHQPSAPAG 154
              |||  |  ::         : ||| :  |: ||   :|     | | ||||    :
Db       1086 IGLNDRTVERD--------FQWTDNTGLQYENWREKQPDNFFAGGEDCVVMVAHENG--- 1134

Qy        155 IGGPYMFQWNDDRCNMKNNFICK 177
                    :|||  ||    ::||
Db       1135 -------RWNDVPCNYNLPYVCK 1150


RESULT 8
PGCV_RAT
ID   PGCV_RAT       STANDARD;      PRT;  2738 AA.
AC   Q9ERB4; O08592; O88564; Q9R1K4;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Versican core protein precursor (Large fibroblast proteoglycan)
DE   (Chondroitin sulfate proteoglycan core protein 2) (PG-M) (Glial
DE   hyaluronate-binding protein) (GHAP) (Fragments).
GN   CPSG2.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE OF 349-2738 FROM N.A. (ISOFORM V0), SEQUENCE FROM N.A.
RP   (ISOFORM V3), AND SEQUENCE OF 2657-2738 FROM N.A. (ISOFORM VINT).
RC   STRAIN=Wistar Kyoto;
RX   MEDLINE=99327053; PubMed=10397680;
RA   Lemire J.M., Braun K.R., Maurel P., Kaplan E.D., Schwartz S.M.,
RA   Wight T.N.;
RT   "Versican/PG-M isoforms in vascular smooth muscle cells.";
RL   Arterioscler. Thromb. Vasc. Biol. 19:1630-1639(1999).
RN   [2]
RP   SEQUENCE OF 349-2738 FROM N.A. (ISOFORM V0).
RC   STRAIN=Wistar Kyoto;
RX   MEDLINE=98308094; PubMed=9642104;
RA   Milev P., Maurel P., Chiba A., Mevissen M., Popp S., Yamaguchi Y.,
```

```
RA   Margolis R.K., Margolis R.U.;
RT   "Differential regulation of expression of hyaluronan-binding
RT   proteoglycans in developing brain: aggrecan, versican, neurocan, and
RT   brevican.";
RL   Biochem. Biophys. Res. Commun. 247:207-212(1998).
RN   [3]
RP   SEQUENCE OF 2421-2463 FROM N.A. (ISOFORM V0).
RC   TISSUE=Kidney;
RX   MEDLINE=98094159; PubMed=9434070;
RA   Pyke C., Kristensen P., Ostergaard P.B., Oturai P.S., Romer J.;
RT   "Proteoglycan expression in the normal rat kidney.";
RL   Nephron 77:461-470(1997).
RN   [4]
RP   SEQUENCE OF 2535-2738 FROM N.A.
RC   STRAIN=Sprague-Dawley; TISSUE=Lung;
RA   Blomberg L.A., Chan W.-Y., Clerch L., Massaro D.;
RT   "Molecular cloning and characterization of two developmentally
RT   regulated genes in rat lung.";
RL   Submitted (SEP-2000) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: May play a role in intercellular signaling and in
CC       connecting cells with the extracellular matrix. May take part in
CC       the regulation of cell motility, growth and differentiation. Binds
CC       hyaluronic acid.
CC   -!- SUBUNIT: Interacts with FBLN1 (By similarity).
CC   -!- SUBCELLULAR LOCATION: Secreted; extracellular matrix.
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=3;
CC         Comment=Additional isoforms seem to exist;
CC       Name=V0;
CC         IsoId=Q9ERB4-1; Sequence=Displayed;
CC       Name=V3;
CC         IsoId=Q9ERB4-2; Sequence=VSP_003091;
CC       Name=Vint;
CC         IsoId=Q9ERB4-3; Sequence=VSP_003092;
CC   -!- TISSUE SPECIFICITY: In kidney is expressed in the papillary area,
CC       but not in glomeruli.
CC   -!- DEVELOPMENTAL STAGE: Disappears after the cartilage development
CC       (By similarity).
CC   -!- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
CC   -!- SIMILARITY: Contains 2 link domains.
CC   -!- SIMILARITY: Contains 2 EGF-like domains.
CC   -!- SIMILARITY: Contains 1 C-type lectin family domain.
CC   -!- SIMILARITY: Contains 1 Sushi (SCR) domain.
CC   -!- SIMILARITY: Belongs to the aggrecan/versican proteoglycan family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF062402; AAC40166.1; -.
DR   EMBL; U75306; AAB51125.1; -.
DR   EMBL; AF084544; AAD48544.1; -.
DR   EMBL; AF072892; AAC26116.1; -.
DR   EMBL; AY007691; AAG16631.1; -.
DR   HSSP; P01132; 1EPG.
DR   InterPro; IPR000152; Asx_hydroxyl_S.
DR   InterPro; IPR000742; EGF_2.
DR   InterPro; IPR001881; EGF_Ca.
DR   InterPro; IPR006209; EGF_like.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003599; Ig.
DR   InterPro; IPR001304; Lectin_C.
DR   InterPro; IPR000538; Link.
DR   InterPro; IPR000436; Sushi_SCR_CCP.
DR   PRINTS; PR01265; LINKMODULE.
DR   ProDom; PD000918; Link; 2.
DR   SMART; SM00032; CCP; 1.
DR   SMART; SM00034; CLECT; 1.
DR   SMART; SM00179; EGF_CA; 1.
DR   SMART; SM00409; IG; 1.
DR   SMART; SM00445; LINK; 2.
DR   PROSITE; PS00010; ASX_HYDROXYL; 1.
DR   PROSITE; PS00615; C_TYPE_LECTIN_1; 1.
DR   PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
DR   PROSITE; PS00022; EGF_1; 2.
DR   PROSITE; PS01186; EGF_2; 1.
DR   PROSITE; PS50026; EGF_3; 2.
DR   PROSITE; PS01187; EGF_CA; 1.
DR   PROSITE; PS50835; IG_LIKE; 1.
DR   PROSITE; PS01241; LINK; 2.
KW   Glycoprotein; Proteoglycan; Lectin; Extracellular matrix; Sushi;
KW   Signal; Repeat; EGF-like domain; Calcium; Immunoglobulin domain;
KW   Hyaluronic acid; Alternative splicing.
FT   SIGNAL        1     20       POTENTIAL.
FT   CHAIN        21   2738       VERSICAN CORE PROTEIN.
FT   NON_CONS    348    349
FT   DOMAIN       21    146       IG-LIKE V-TYPE.
FT   DOMAIN      167    244       LINK 1.
FT   DOMAIN      265    346       LINK 2.
FT   DOMAIN     <349    695       GAG-ALPHA
FT                                (GLUCOSAMINOGLYCAN ATTACHMENT DOMAIN).
FT   DOMAIN      696   2431       GAG-BETA.
FT   DOMAIN     2431   2467       EGF-LIKE 1.
FT   DOMAIN     2469   2505       EGF-LIKE 2, CALCIUM-BINDING.
FT   DOMAIN     2518   2632       C-TYPE LECTIN.
FT   DOMAIN     2637   2695       SUSHI.
FT   DISULFID     44    130       BY SIMILARITY.
FT   DISULFID    172    243       BY SIMILARITY.
FT   DISULFID    196    217       BY SIMILARITY.
FT   DISULFID    270    345       BY SIMILARITY.
FT   DISULFID    294    315       BY SIMILARITY.
FT   DISULFID   2435   2446       BY SIMILARITY.
FT   DISULFID   2440   2455       BY SIMILARITY.
FT   DISULFID   2457   2466       BY SIMILARITY.
FT   DISULFID   2473   2484       BY SIMILARITY.
FT   DISULFID   2478   2493       BY SIMILARITY.
FT   DISULFID   2495   2504       BY SIMILARITY.
FT   DISULFID   2511   2522       BY SIMILARITY.
FT   DISULFID   2539   2631       BY SIMILARITY.
FT   DISULFID   2607   2623       BY SIMILARITY.
FT   DISULFID   2638   2681       BY SIMILARITY.
FT   DISULFID   2667   2694       BY SIMILARITY.
FT   CARBOHYD     57     57       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    330    330       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    692    692       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    758    758       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    802    802       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    805    805       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1257   1257       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1435   1435       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1633   1633       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1660   1660       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1684   1684       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1738   1738       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1848   1848       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   2004   2004       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   2409   2409       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   2711   2711       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   2721   2721       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   VARSPLIC    349   2431       Missing (in isoform V3).
FT                                /FTId=VSP_003091.
FT   VARSPLIC   2697   2738       PSAYQRTYSKKYLKNSSSVKDNSINTSKHEHRWSRRWQETR
FT                                R -> RKWSFRKNGQPCFNKY (in isoform Vint).
FT                                /FTId=VSP_003092.
FT   CONFLICT   2535   2539       AEREC -> NSARG (IN REF. 4).
SQ   SEQUENCE   2738 AA;  300004 MW;  12CA626D58BD8C6A CRC64;

  Query Match             8.7%;  Score 174;  DB 1;  Length 2738;
  Best Local Similarity   28.5%;  Pred. No. 5.8e-06;
  Matches   47;  Conservative   23;  Mismatches   55;  Indels   40;  Gaps    8;

Qy         38 QRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDF- 96
```

```
Db       2519 QGQCYK--YF---AHRRTWDAAERECRLQGAHLTSILSHEEQMFVNRV-------GHDYQ 2566

Qy         97 WIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEP----SCGSEVCVVMYHQPSAP 152

Db       2567 WIGL--------NDKMFEHDFRWTDGSALQYENWRPNQPDSFFSAGEDCVVIIWHENG-- 2616

Qy        153 AGIGGPYMFQWNDDRCNMKNNFICKYS----DEKPAVPSREAEGE 193

Db       2617 ---------QWNDVPCNYHLTYTCKKGTVACGQPPVVENAKTFGK 2652


RESULT 9
PGCV_MOUSE
ID   PGCV_MOUSE     STANDARD;      PRT;  3358 AA.
AC   Q62059; Q62058; Q9CUU0;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   Versican core protein precursor (Large fibroblast proteoglycan)
DE   (Chondroitin sulfate proteoglycan core protein 2) (PG-M).
GN   CSPG2.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORMS V0; V1 AND V2).
RC   STRAIN=C57BL/6, and Swiss Webster; TISSUE=Brain;
RX   MEDLINE=95122551; PubMed=7822336;
RA   Ito K., Shinomura T., Zako M., Ujita M., Kimata K.;
RT   "Multiple forms of mouse PG-M, a large chondroitin sulfate
RT   proteoglycan generated by alternative splicing.";
RL   J. Biol. Chem. 270:958-965(1995).
RN   [2]
RP   SEQUENCE FROM N.A. (ISOFORM V3).
RC   STRAIN=C57BL/6;
RX   MEDLINE=95181355; PubMed=7876137;
RA   Zako M., Shinomura T., Ujita M., Ito K., Kimata K.;
RT   "Expression of PG-M(V3), an alternatively spliced form of PG-M
RT   without a chondroitin sulfate attachment in region in mouse and human
RT   tissues.";
RL   J. Biol. Chem. 270:3914-3918(1995).
RN   [3]
RP   SEQUENCE OF 1-1692 FROM N.A. (ISOFORM V1).
RC   STRAIN=C57BL/6J; TISSUE=Skin;
RX   MEDLINE=22354683; PubMed=12466851;
RA   Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
RA   Nikaido I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,
RA   Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,
RA   Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
RA   Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
RA   Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.E., Cousins S.,
RA   Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
RA   Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
RA   Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
RA   Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,
RA   Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
RA   Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
RA   Nagashima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G.,
RA   Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
RA   Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
RA   Sandelin A., Schneider C., Semple C.A., Setou M., Shimada K.,
RA   Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
RA   Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,
RA   Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L.,
RA   Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
RA   Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,
RA   Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
RA   Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
RA   Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
RA   Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
RA   Birney E., Hayashizaki Y.;
RT   "Analysis of the mouse transcriptome based on functional annotation of
RT   60,770 full-length cDNAs.";
RL   Nature 420:563-573(2002).
RN   [4]
RP   INTERACTION WITH FBLN1.
RX   MEDLINE=99329059; PubMed=10400671;
RA   Aspberg A., Adam S., Kostka G., Timpl R., Heinegaard D.;
RT   "Fibulin-1 is a ligand for the C-type lectin domains of aggrecan and
RT   versican.";
RL   J. Biol. Chem. 274:20444-20449(1999).
CC   -!- FUNCTION: May play a role in intercellular signaling and in
CC       connecting cells with the extracellular matrix. May take part in
CC       the regulation of cell motility, growth and differentiation. Binds
CC       hyaluronic acid.
CC   -!- SUBUNIT: Interacts with FBLN1.
CC   -!- SUBCELLULAR LOCATION: Secreted; extracellular matrix.
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=4;
CC         Comment=Additional isoforms seem to exist;
CC       Name=V0;
CC         IsoId=Q62059-1; Sequence=Displayed;
CC       Name=V1;
CC         IsoId=Q62059-2; Sequence=VSP_003087, VSP_003088;
CC       Name=V2;
CC         IsoId=Q62059-3; Sequence=VSP_003089;
CC       Name=V3;
CC         IsoId=Q62059-4; Sequence=VSP_003087, VSP_003090;
CC   -!- TISSUE SPECIFICITY: V2 is found only in brain.
CC   -!- DEVELOPMENTAL STAGE: Disappears after the cartilage development.
CC   -!- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
CC   -!- SIMILARITY: Contains 2 link domains.
CC   -!- SIMILARITY: Contains 2 EGF-like domains.
CC   -!- SIMILARITY: Contains 1 C-type lectin family domain.
CC   -!- SIMILARITY: Contains 1 Sushi (SCR) domain.
CC   -!- SIMILARITY: Belongs to the aggrecan/versican proteoglycan family.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; D16263; BAA03796.1; -.
DR   EMBL; D28599; -; NOT_ANNOTATED_CDS.
DR   EMBL; D32040; BAA06802.1; -.
DR   EMBL; AK014525; BAB29411.2; -.
DR   HSSP; P01132; 1EPG.
DR   MGD; MGI:102889; Cspg2.
DR   InterPro; IPR000152; Asx_hydroxyl_S.
DR   InterPro; IPR000742; EGF_2.
DR   InterPro; IPR001881; EGF_Ca.
DR   InterPro; IPR006209; EGF_like.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003599; Ig.
DR   InterPro; IPR001304; Lectin_C.
DR   InterPro; IPR000538; Link.
DR   InterPro; IPR000436; Sushi_SCR_CCP.
DR   Pfam; PF00008; EGF; 2.
DR   Pfam; PF00047; ig; 1.
DR   Pfam; PF00059; lectin_c; 1.
DR   Pfam; PF00084; sushi; 1.
DR   Pfam; PF00193; Xlink; 2.
DR   PRINTS; PR01265; LINKMODULE.
DR   ProDom; PD000918; Link; 2.
DR   SMART; SM00032; CCP; 1.
DR   SMART; SM00034; CLECT; 1.
DR   SMART; SM00179; EGF_CA; 1.
DR   SMART; SM00409; IG; 1.
DR   SMART; SM00445; LINK; 2.
DR   PROSITE; PS00010; ASX_HYDROXYL; 1.
DR   PROSITE; PS00615; C_TYPE_LECTIN_1; 1.
```

```
DR   PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
DR   PROSITE; PS00022; EGF_1; 2.
DR   PROSITE; PS01186; EGF_2; 1.
DR   PROSITE; PS50026; EGF_3; 2.
DR   PROSITE; PS01187; EGF_CA; 1.
DR   PROSITE; PS50835; IG_LIKE; 1.
DR   PROSITE; PS01241; LINK; 2.
KW   Glycoprotein; Proteoglycan; Lectin; Extracellular matrix; Sushi;
KW   Signal; Repeat; EGF-like domain; Calcium; Immunoglobulin domain;
KW   Hyaluronic acid; Alternative splicing.
FT   SIGNAL        1     20       POTENTIAL.
FT   CHAIN        21   3358       VERSICAN CORE PROTEIN.
FT   DOMAIN       21    146       IG-LIKE V-TYPE.
FT   DOMAIN      167    244       LINK 1.
FT   DOMAIN      265    346       LINK 2.
FT   DOMAIN      348   1308       GAG-ALPHA
FT                                (GLUCOSAMINOGLYCAN ATTACHMENT DOMAIN).
FT   DOMAIN     1309   3052       GAG-BETA.
FT   DOMAIN     3052   3088       EGF-LIKE 1.
FT   DOMAIN     3090   3126       EGF-LIKE 2, CALCIUM-BINDING (POTENTIAL).
FT   DOMAIN     3139   3253       C-TYPE LECTIN.
FT   DOMAIN     3258   3316       SUSHI.
FT   DISULFID     44    130       BY SIMILARITY.
FT   DISULFID    172    243       BY SIMILARITY.
FT   DISULFID    196    217       BY SIMILARITY.
FT   DISULFID    270    333       BY SIMILARITY.
FT   DISULFID    294    315       BY SIMILARITY.
FT   DISULFID   3056   3067       BY SIMILARITY.
FT   DISULFID   3061   3076       BY SIMILARITY.
FT   DISULFID   3078   3087       BY SIMILARITY.
FT   DISULFID   3094   3105       BY SIMILARITY.
FT   DISULFID   3099   3114       BY SIMILARITY.
FT   DISULFID   3116   3125       BY SIMILARITY.
FT   DISULFID   3132   3143       BY SIMILARITY.
FT   DISULFID   3160   3252       BY SIMILARITY.
FT   DISULFID   3228   3244       BY SIMILARITY.
FT   DISULFID   3259   3302       BY SIMILARITY.
FT   DISULFID   3288   3315       BY SIMILARITY.
FT   CARBOHYD     57     57       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    330    330       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    351    351       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    441    441       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    807    807       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    914    914       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    951    951       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1305   1305       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1372   1372       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1679   1679       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   2054   2054       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   2244   2244       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   2362   2362       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   2627   2627       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   3030   3030       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   3332   3332       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   3342   3342       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   VARSPLIC    348    348       P -> R (in isoform V1 and isoform V3).
FT                                /FTId=VSP_003087.
FT   VARSPLIC    349   1308       Missing (in isoform V1).
FT                                /FTId=VSP_003088.
FT   VARSPLIC   1309   3052       Missing (in isoform V2).
FT                                /FTId=VSP_003089.
FT   VARSPLIC    349   3052       Missing (in isoform V3).
FT                                /FTId=VSP_003090.
FT   CONFLICT    126    126       A -> G (IN REF. 3).
FT   CONFLICT    348    348       MISSING (IN REF. 3).
FT   CONFLICT   1658   1658       I -> T (IN REF. 3).
FT   CONFLICT   1674   1680       TVWNSNS -> QFGIQTA (IN REF. 3).
SQ   SEQUENCE   3358 AA;  366938 MW;  071B80026BC0762D CRC64;

  Query Match             8.7%;  Score 174;  DB 1;  Length 3358;
  Best Local Similarity  28.5%;  Pred. No. 7.6e-06;
  Matches   47;  Conservative   23;  Mismatches   55;  Indels   40;  Gaps    8;

Qy        38 QRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDF- 96
             |  |||  ||   : |  :: |:  ||  |  | || | :||  : :          |:
Db      3140 QGQCYK--YF---AHRRTWDAAERECRLQGAHLTSILSHEEQMFVNRV-------GHDYQ 3187

Qy        97 WIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEP----SCGSEVCVVMYHQPSAP 152
             ||||         |    :  : |||||  |: ||  ::|    | | :  |:::|:
Db      3188 WIGL--------NDKMFEHDFRWTDGSALQYENWRPNQPDSFFSAGEDCVVIIWHENG-- 3237

Qy       153 AGIGGPYMFQWNDDRCNMKNNFICKYS----DEKPAVPSREAEGE 193
                 ||||  ||    : ||       : | | : :  |:
Db      3238 ---------QWNDVPCNYHLTYTCKKGTVACGQPPVVENAKTFGK 3273


RESULT 10
PGCV_HUMAN
ID   PGCV_HUMAN     STANDARD;      PRT;  3396 AA.
AC   P13611; P20754; Q13010; Q13189; Q15123; Q9UNW5;
DT   01-JAN-1990 (Rel. 13, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Versican core protein precursor (Large fibroblast proteoglycan)
DE   (Chondroitin sulfate proteoglycan core protein 2) (PG-M) (Glial
DE   hyaluronate-binding protein) (GHAP).
GN   CSPG2.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORM V0).
RX   MEDLINE=95105188; PubMed=7528742;
RA   Naso M.F., Zimmermann D.R., Iozzo R.V.;
RT   "Characterization of the complete genomic structure of the human
RT   versican gene and functional analysis of its promoter.";
RL   J. Biol. Chem. 269:32999-33008(1994).
RN   [2]
RP   SEQUENCE FROM N.A. (ISOFORM V1).
RC   TISSUE=Placenta;
RX   MEDLINE=90059882; PubMed=2583089;
RA   Zimmermann D.R., Ruoslahti E.;
RT   "Multiple domains of the large fibroblast proteoglycan, versican.";
RL   EMBO J. 8:2975-2981(1989).
RN   [3]
RP   SEQUENCE FROM N.A. (ISOFORM V2).
RC   TISSUE=Glial tumor;
RX   MEDLINE=95105187; PubMed=7806529;
RA   Dours-Zimmermann M.T., Zimmermann D.R.;
RT   "A novel glycosaminoglycan attachment domain identified in two
RT   alternative splice variants of human versican.";
RL   J. Biol. Chem. 269:32992-32998(1994).
RN   [4]
RP   SEQUENCE OF 2711-3396 FROM N.A.
RC   TISSUE=Lung fibroblast;
RX   MEDLINE=88007514; PubMed=2820964;
RA   Krusius T., Gehlsen K.R., Ruoslahti E.;
RT   "A fibroblast chondroitin sulfate proteoglycan core protein contains
RT   lectin-like and growth factor-like sequences.";
RL   J. Biol. Chem. 262:13120-13125(1987).
RN   [5]
RP   SEQUENCE OF 251-347 FROM N.A.
RX   MEDLINE=93122792; PubMed=1478664;
RA   Iozzo R.V., Naso M.F., Cannizzaro L.A., Wasmuth J.J.,
RA   McPherson J.D.;
RT   "Mapping of the versican proteoglycan gene (CSPG2) to the long arm of
RT   human chromosome 5 (5q12-5q14).";
RL   Genomics 14:845-851(1992).
RN   [6]
RP   SEQUENCE FROM N.A. (ISOFORM V3).
RC   TISSUE=Brain;
RX   MEDLINE=95181355; PubMed=7876137;
RA   Zako M., Shinomura T., Ujita M., Ito K., Kimata K.;
RT   "Expression of PG-M(V3), an alternatively spliced form of PG-M
RT   without a chondroitin sulfate attachment in region in mouse and human
```

```
RT   tissues.";
RL   J. Biol. Chem. 270:3914-3918(1995).
RN   [7]
RP   SEQUENCE OF 3333-3396 FROM N.A. (ISOFORM VINT).
RC   TISSUE=Aortic smooth muscle;
RX   MEDLINE=99327053; PubMed=10397680;
RA   Lemire J.M., Braun K.R., Maurel P., Kaplan E.D., Schwartz S.M.,
RA   Wight T.N.;
RT   "Versican/PG-M isoforms in vascular smooth muscle cells.";
RL   Arterioscler. Thromb. Vasc. Biol. 19:1630-1639(1999).
RN   [8]
RP   PARTIAL SEQUENCE.
RC   TISSUE=Brain;
RX   MEDLINE=89174663; PubMed=2466833;
RA   Perides G., Lane W.S., Andrews D., Dahl D., Bignami A.;
RT   "Isolation and partial characterization of a glial
RT   hyaluronate-binding protein.";
RL   J. Biol. Chem. 264:5981-5987(1989).
RN   [9]
RP   TISSUE SPECIFICITY OF ISOFORMS.
RX   MEDLINE=96213482; PubMed=8627343;
RA   Paulus W., Baur I., Dours-Zimmermann M.T., Zimmermann D.R.;
RT   "Differential expression of versican isoforms in brain tumors.";
RL   J. Neuropathol. Exp. Neurol. 55:528-533(1996).
CC   -!- FUNCTION: May play a role in intercellular signaling and in
CC       connecting cells with the extracellular matrix. May take part in
CC       the regulation of cell motility, growth and differentiation. Binds
CC       hyaluronic acid.
CC   -!- SUBUNIT: Interacts with FBLN1 (By similarity).
CC   -!- SUBCELLULAR LOCATION: Secreted; extracellular matrix.
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=5;
CC         Comment=Additional isoforms seem to exist;
CC       Name=V0;
CC         IsoId=P13611-1; Sequence=Displayed;
CC       Name=V1;
CC         IsoId=P13611-2; Sequence=VSP_003082, VSP_003083;
CC       Name=V2;
CC         IsoId=P13611-3; Sequence=VSP_003084;
CC       Name=V3;
CC         IsoId=P13611-4; Sequence=VSP_003082, VSP_003085;
CC       Name=Vint;
CC         IsoId=P13611-5; Sequence=VSP_003086;
CC   -!- TISSUE SPECIFICITY: Cerebral white matter. V0 and V1 is expressed
CC       in normal brain, gliomas, medulloblastomas, schwannomas,
CC       neurofibromas, and menningiomas; v2 is restricted to normal brain
CC       and gliomas; v3 is found in all these tissues except
CC       medulloblastomas.
CC   -!- DEVELOPMENTAL STAGE: Disappears after the cartilage development.
CC   -!- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
CC   -!- SIMILARITY: Contains 2 link domains.
CC   -!- SIMILARITY: Contains 2 EGF-like domains.
CC   -!- SIMILARITY: Contains 1 C-type lectin family domain.
CC   -!- SIMILARITY: Contains 1 Sushi (SCR) domain.
CC   -!- SIMILARITY: Belongs to the aggrecan/versican proteoglycan family.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; U16306; AAA65018.1; -.
DR   EMBL; X15998; CAA34128.1; -.
DR   EMBL; S52488; AAB24878.1; -.
DR   EMBL; U26555; AAA67565.1; -.
DR   EMBL; D32039; BAA06801.1; -.
DR   EMBL; J02814; AAA36437.1; -.
DR   EMBL; AF084545; AAD48545.1; -.
DR   PIR; S06014; A60979.
DR   HSSP; P01132; 1EGF.
DR   Genew; HGNC:2464; CSPG2.
DR   MIM; 118661; -.
DR   GO; GO:0005578; C:extracellular matrix; TAS.
DR   GO; GO:0005540; F:hyaluronic acid binding; TAS.
DR   GO; GO:0008037; P:cell recognition; TAS.
DR   GO; GO:0007275; P:development; TAS.
DR   InterPro; IPR000152; Asx_hydroxyl_S.
DR   InterPro; IPR000742; EGF_2.
DR   InterPro; IPR001881; EGF_Ca.
DR   InterPro; IPR006209; EGF_like.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003599; Ig.
DR   InterPro; IPR001304; Lectin_C.
DR   InterPro; IPR000538; Link.
DR   InterPro; IPR000436; Sushi_SCR_CCP.
DR   Pfam; PF00008; EGF; 2.
DR   Pfam; PF00059; lectin_c; 1.
DR   Pfam; PF00084; sushi; 1.
DR   Pfam; PF00193; Xlink; 2.
DR   PRINTS; PR01265; LINKMODULE.
DR   ProDom; PD000918; Link; 2.
DR   SMART; SM00032; CCP; 1.
DR   SMART; SM00034; CLECT; 1.
DR   SMART; SM00179; EGF_CA; 1.
DR   SMART; SM00409; IG; 1.
DR   SMART; SM00445; LINK; 2.
DR   PROSITE; PS00010; ASX_HYDROXYL; 1.
DR   PROSITE; PS00615; C_TYPE_LECTIN_1; 1.
DR   PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
DR   PROSITE; PS00022; EGF_1; 2.
DR   PROSITE; PS01186; EGF_2; 1.
DR   PROSITE; PS50026; EGF_3; 2.
DR   PROSITE; PS01187; EGF_CA; 1.
DR   PROSITE; PS50835; IG_LIKE; 1.
DR   PROSITE; PS01241; LINK; 2.
KW   Glycoprotein; Proteoglycan; Lectin; Extracellular matrix; Sushi;
KW   Signal; Repeat; EGF-like domain; Calcium; Immunoglobulin domain;
KW   Hyaluronic acid; Alternative splicing.
FT   SIGNAL        1     20       POTENTIAL.
FT   CHAIN        21   3396       VERSICAN CORE PROTEIN.
FT   DOMAIN       21    146       IG-LIKE V-TYPE.
FT   DOMAIN      167    244       LINK 1.
FT   DOMAIN      265    346       LINK 2.
FT   DOMAIN      348   1335       GAG-ALPHA
FT                                (GLUCOSAMINOGLYCAN ATTACHMENT DOMAIN).
FT   DOMAIN     1336   3089       GAG-BETA.
FT   DOMAIN     3089   3125       EGF-LIKE 1.
FT   DOMAIN     3127   3163       EGF-LIKE 2, CALCIUM-BINDING (POTENTIAL).
FT   DOMAIN     3176   3290       C-TYPE LECTIN.
FT   DOMAIN     3295   3353       SUSHI.
FT   DISULFID     44    130       BY SIMILARITY.
FT   DISULFID    172    243       BY SIMILARITY.
FT   DISULFID    196    217       BY SIMILARITY.
FT   DISULFID    270    345       BY SIMILARITY.
FT   DISULFID    294    315       BY SIMILARITY.
FT   DISULFID   3093   3104       BY SIMILARITY.
FT   DISULFID   3098   3113       BY SIMILARITY.
FT   DISULFID   3115   3124       BY SIMILARITY.
FT   DISULFID   3131   3142       BY SIMILARITY.
FT   DISULFID   3136   3151       BY SIMILARITY.
FT   DISULFID   3153   3162       BY SIMILARITY.
FT   DISULFID   3169   3180       BY SIMILARITY.
FT   DISULFID   3197   3289       BY SIMILARITY.
FT   DISULFID   3265   3281       BY SIMILARITY.
FT   DISULFID   3296   3339       BY SIMILARITY.
FT   DISULFID   3325   3352       BY SIMILARITY.
FT   CARBOHYD     57     57       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    330    330       N-LINKED (GLCNAC. . .).
FT   CARBOHYD    615    615       N-LINKED (GLCNAC. . .) (POTENTIAL).

  Query Match             8.7%;  Score 174;  DB 1;  Length 3396;
  Best Local Similarity   28.5%;  Pred. No. 7.7e-06;
  Matches   47;  Conservative   23;  Mismatches   55;  Indels   40;  Gaps    8;
```

```
Qy         38 QRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDF- 96
              |  |||  ||   : |  :: |:  ||  |  | || | :||  : :          |:
Db       3177 QGQCYK--YF---AHRRTWDAAERECRLQGAHLTSILSHEEQMFVNRV-------GHDYQ 3224

Qy         97 WIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEP----SCGSEVCVVMYHQPSAP 152
              ||||        |    :  : |||||  |: ||  ::|    | | :  |:::|:
Db       3225 WIGL--------NDKMFEHDFRWTDGSTLQYENWRPNQPDSFFSAGEDCVVIIWHENG-- 3274

Qy        153 AGIGGPYMFQWNDDRCNMKNNFICKYS----DEKPAVPSREAEGE 193
                     ||||  ||    : ||       : | | : :  |:
Db       3275 ---------QWNDVPCNYHLTYTCKKGTVACGQPPVVENAKTFGK 3310
```

RESULT 11
CD93_RAT
ID CD93_RAT STANDARD; PRT; 643 AA.
AC Q9ET61; Q9JIZ6;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Complement component C1q receptor precursor (Complement component 1, q
DE subcomponent, receptor 1) (C1qRp) (C1qR(p)) (C1q/MBL/SPA receptor)
DE (CD93 antigen) (Cell surface antigen AA4).
GN C1QR1 OR CD93 OR C1QRP.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PVG; TISSUE=Natural killer cells;
RX MEDLINE=20545218; PubMed=11093152;
RA Lovik G., Vaage J.T., Dissen E., Szpirer C., Ryan J.C., Rolstad B.;
RT "Characterization and molecular cloning of rat C1qRp, a receptor on NK
RT cells.";
RL Eur. J. Immunol. 30:3355-3362(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Wistar; TISSUE=Lung;
RX MEDLINE=20507883; PubMed=10934210;
RA Dean Y.D., McGreal E.P., Akatsu H., Gasque P.;
RT "Molecular and cellular properties of the rat AA4 antigen, a C-type
RT lectin-like receptor with structural homology to thrombomodulin.";
RL J. Biol. Chem. 275:34382-34392(2000).
CC -!- FUNCTION: Receptor (or element of a larger receptor complex) for
CC   C1q, mannose-binding lectin (MBL2) and pulmonary surfactant
CC   protein A (SPA). May mediate the enhancement of phagocytosis in
CC   monocytes and macrophages upon interaction with soluble defense
CC   collagens. May play a role in intercellular adhesion.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- TISSUE SPECIFICITY: Widely expressed. Highly expressed in lung and
CC   heart. Expressed at lower level in brain, thymus, liver, spleen,
CC   intestine, kidney, adrenal gland, muscle and testis. Expressed on
CC   endothelial cells, platelets, undifferentiated monocytes and
CC   circulating natural killer cells.
CC -!- PTM: N- and O-glycosylated (By similarity).
CC -!- SIMILARITY: Contains 1 C-type lectin family domain.
CC -!- SIMILARITY: Contains 5 EGF-like domains.
CC --
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC --
DR EMBL; AF136537; AAG01572.1; -.
DR EMBL; AF160978; AAF80402.1; -.
DR HSSP; P35555; 1EMN.
DR GO; GO:0016021; C:integral to membrane; ISS.
DR GO; GO:0004872; F:receptor activity; ISS.
DR GO; GO:0016337; P:cell-cell adhesion; ISS.
DR GO; GO:0042116; P:macrophage activation; ISS.
DR GO; GO:0006909; P:phagocytosis; ISS.
DR InterPro; IPR000152; Asx_hydroxyl_S.
DR InterPro; IPR001881; EGF_Ca.
DR InterPro; IPR006209; EGF_like.
DR InterPro; IPR001304; Lectin_C.
DR Pfam; PF00008; EGF; 4.
DR Pfam; PF00059; lectin_c; 1.
DR SMART; SM00034; CLECT; 1.
DR SMART; SM00179; EGF_CA; 3.
DR PROSITE; PS00010; ASX_HYDROXYL; 3.
DR PROSITE; PS00615; C_TYPE_LECTIN_1; FALSE_NEG.
DR PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
DR PROSITE; PS01186; EGF_2; 3.
DR PROSITE; PS50026; EGF_3; 4.
DR PROSITE; PS01187; EGF_CA; 3.
KW Cell adhesion; Receptor; Repeat; Signal; Transmembrane;
KW EGF-like domain; Lectin; Glycoprotein.
FT SIGNAL   1   23   POTENTIAL.
FT CHAIN   24   643   COMPLEMENT COMPONENT C1Q RECEPTOR.
FT DOMAIN   24   571   EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 572   592   POTENTIAL.
FT DOMAIN   593   643   CYTOPLASMIC (POTENTIAL).
FT DOMAIN   31   173   C-TYPE LECTIN.
FT DOMAIN   257   298   EGF-LIKE 1.
FT DOMAIN   299   341   EGF-LIKE 2.
FT DOMAIN   342   381   EGF-LIKE 3, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN   382   423   EGF-LIKE 4, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN   424   462   EGF-LIKE 5, CALCIUM-BINDING (POTENTIAL).
FT DISULFID 261   272   BY SIMILARITY.
FT DISULFID 268   282   BY SIMILARITY.
FT DISULFID 284   297   BY SIMILARITY.
FT DISULFID 303   314   BY SIMILARITY.
FT DISULFID 308   325   BY SIMILARITY.
FT DISULFID 327   340   BY SIMILARITY.
FT DISULFID 346   355   BY SIMILARITY.
FT DISULFID 351   364   BY SIMILARITY.
FT DISULFID 366   380   BY SIMILARITY.
FT DISULFID 386   397   BY SIMILARITY.
FT DISULFID 393   406   BY SIMILARITY.
FT DISULFID 408   422   BY SIMILARITY.
FT DISULFID 428   437   BY SIMILARITY.
FT DISULFID 433   446   BY SIMILARITY.
FT DISULFID 448   461   BY SIMILARITY.
FT CARBOHYD 322   322   N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 498   498   N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 417   417   E -> K (IN REF. 2).
SQ SEQUENCE 643 AA; 68781 MW; 9AE4C933AD943DB6 CRC64;

  Query Match 8.6%; Score 172.5; DB 1; Length 643;
  Best Local Similarity 25.9%; Pred. No. 1.1e-06;
  Matches 56; Conservative 37; Mismatches 74; Indels 49; Gaps 11;

```
Qy         14 VLLVGLRAATGRLLSGQP------VCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDG 67
              :||:||    |:| :|       || |     ||   :      :|:  ||:  |  :|
Db          8 LLLLGL---LGQLWAGAAADSEAVVCEG---TACYTAHW-----GKLSAAEAQHRCNENG 56

Qy         68 GQLVSIESEDEQKLIEKFIENLL----PSD---GDFWIGLRRREEKQSNSTACQDLYAWT 120
              | | :::||:| : ::: :  ||    ||:   | |||||:| : | :        ::|
Db         57 GNLATVKSEEEARHVQEALAQLLKTKAPSETKIGKFWIGLQREKGKCTYHDLPMKGFSWV 116

Qy        121 -DGSISQFRNWY-VDEPSCGSEVCVVMY-------HQPSAPAGIGGPYMFQWNDDRC--- 168
               |  : : |||   : || |: || :        |     |          :|::  |
Db        117 GGGEDTTYSNWYKASKSSCISKRCVSLILDLSLKPHPSHLP---------KWHESPCGTP 167

Qy        169 ----NMKNNFICKYSDEKPAVPSREAEGEETELTTP 200
                  |    |:||:: :    |       :   |||
Db        168 DAPGNSIEGFLCKFNFKGMCSPLALGGPGQLTYTTP 203
```

RESULT 12
PGCV_CHICK

```
ID   PGCV_CHICK     STANDARD;      PRT;  3562 AA.
AC   Q90953; Q90945;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Versican core protein precursor (Large fibroblast proteoglycan)
DE   (Chondroitin sulfate proteoglycan core protein 2) (PG-M).
GN   CPSG2.
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
OX   NCBI_TaxID=9031;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORMS V0 AND V1).
RC   STRAIN=White leghorn; TISSUE=Limb bud;
RX   MEDLINE=93300846; PubMed=8314802;
RA   Shinomura T., Nishida Y., Ito K., Kimata K.;
RT   "cDNA cloning of PG-M, a large chondroitin sulfate proteoglycan
RT   expressed during chondrogenesis in chick limb buds. Alternative
RT   spliced multiforms of PG-M and their relationships to versican.";
RL   J. Biol. Chem. 268:14461-14469(1993).
CC   -!- FUNCTION: May play a role in intercellular signaling and in
CC       connecting cells with the extracellular matrix. May take part in
CC       the regulation of cell motility, growth and differentiation. Binds
CC       hyaluronic acid.
CC   -!- SUBCELLULAR LOCATION: Secreted; extracellular matrix.
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=2;
CC         Comment=Additional isoforms seem to exist;
CC       Name=V0;
CC         IsoId=Q90953-1; Sequence=Displayed;
CC       Name=V1;
CC         IsoId=Q90953-2; Sequence=VSP_003093;
CC   -!- TISSUE SPECIFICITY: Prechondrogenic condensation area of
CC       developing limb buds.
CC   -!- DEVELOPMENTAL STAGE: Disappears after the cartilage development
CC       (By similarity).
CC   -!- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
CC   -!- SIMILARITY: Contains 2 link domains.
CC   -!- SIMILARITY: Contains 2 EGF-like domains.
CC   -!- SIMILARITY: Contains 1 C-type lectin family domain.
CC   -!- SIMILARITY: Contains 1 Sushi (SCR) domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X60226; CAA42787.1; -.
DR   EMBL; D13542; BAA02742.1; -.
DR   PIR; A47171; A47171.
DR   HSSP; P00740; 1EDM.
DR   InterPro; IPR000152; Asx_hydroxyl_S.
DR   InterPro; IPR000742; EGF_2.
DR   InterPro; IPR001881; EGF_Ca.
DR   InterPro; IPR006209; EGF_like.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003599; Ig.
DR   InterPro; IPR001304; Lectin_C.
DR   InterPro; IPR000538; Link.
DR   InterPro; IPR000436; Sushi_SCR_CCP.
DR   Pfam; PF00008; EGF; 2.
DR   Pfam; PF00047; ig; 1.
DR   Pfam; PF00059; lectin_c; 1.
DR   Pfam; PF00084; sushi; 1.
DR   Pfam; PF00193; Xlink; 2.
DR   PRINTS; PR01265; LINKMODULE.
DR   ProDom; PD000918; Link; 2.
DR   SMART; SM00032; CCP; 1.
DR   SMART; SM00034; CLECT; 1.
DR   SMART; SM00179; EGF_CA; 1.
DR   SMART; SM00409; IG; 1.
DR   SMART; SM00445; LINK; 2.
DR   PROSITE; PS00010; ASX_HYDROXYL; 1.
DR   PROSITE; PS00615; C_TYPE_LECTIN_1; 1.
DR   PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
DR   PROSITE; PS00022; EGF_1; 2.
DR   PROSITE; PS01186; EGF_2; 1.
DR   PROSITE; PS50026; EGF_3; 2.
DR   PROSITE; PS01187; EGF_CA; 1.
DR   PROSITE; PS50835; IG_LIKE; 1.
DR   PROSITE; PS01241; LINK; 2.
KW   Glycoprotein; Proteoglycan; Lectin; Extracellular matrix; Sushi;
KW   Signal; Repeat; EGF-like domain; Calcium; Immunoglobulin domain;
KW   Hyaluronic acid; Alternative splicing.
FT   SIGNAL        1     26       POTENTIAL.
FT   CHAIN        27   3562       VERSICAN CORE PROTEIN.
FT   DOMAIN       27    143       IG-LIKE V-TYPE.
FT   DOMAIN      166    243       LINK 1.
FT   DOMAIN      264    345       LINK 2.
FT   DOMAIN     3254   3290       EGF-LIKE 1.
FT   DOMAIN     3292   3328       EGF-LIKE 2, CALCIUM-BINDING (POTENTIAL).
FT   DOMAIN     3341   3455       C-TYPE LECTIN.
FT   DOMAIN     3460   3518       SUSHI.
FT   DISULFID     44    129       BY SIMILARITY.
FT   DISULFID    171    242       BY SIMILARITY.
FT   DISULFID    195    216       BY SIMILARITY.
FT   DISULFID    269    344       BY SIMILARITY.
FT   DISULFID    293    314       BY SIMILARITY.
FT   DISULFID   3258   3269       BY SIMILARITY.
FT   DISULFID   3263   3278       BY SIMILARITY.
FT   DISULFID   3280   3289       BY SIMILARITY.
FT   DISULFID   3296   3307       BY SIMILARITY.
FT   DISULFID   3301   3316       BY SIMILARITY.
FT   DISULFID   3318   3327       BY SIMILARITY.
FT   DISULFID   3334   3345       BY SIMILARITY.
FT   DISULFID   3362   3454       BY SIMILARITY.
FT   DISULFID   3430   3446       BY SIMILARITY.
FT   DISULFID   3461   3504       BY SIMILARITY.
FT   DISULFID   3490   3517       BY SIMILARITY.
FT   CARBOHYD    163    163       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    235    235       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    329    329       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    529    529       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    709    709       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    948    948       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1409   1409       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1479   1479       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1523   1523       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1530   1530       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1625   1625       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1751   1751       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1988   1988       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   2088   2088       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   2089   2089       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   2507   2507       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   2642   2642       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   2679   2679       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   2748   2748       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   2762   2762       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   3069   3069       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   3194   3194       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   3232   3232       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   3545   3545       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   VARSPLIC    485   1411       Missing (in isoform V1).
FT                                /FTId=VSP_003093.
SQ   SEQUENCE   3562 AA;  388078 MW;  9BC566E88C1602D2 CRC64;

  Query Match             8.6%;  Score 171;  DB 1;  Length 3562;
  Best Local Similarity   28.5%;  Pred. No. 1.5e-05;
  Matches   47;  Conservative   23;  Mismatches   55;  Indels   40;  Gaps    8;
```

```
Qy         38 QRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDF- 96
              |  |||  ||   : |  :: |:  ||  |  | || | :||  : :          |:
Db       3342 QGQCYK--YF---AHRRTWDTAERECRLQGAHLTSILSHEEQVFVNRI-------GHDYQ 3389

Qy         97 WIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEP----SCGSEVCVVMYHQPSAP 152
              ||||        |    :  : |||||  |: ||  ::|    | | :  |:::|:
Db       3390 WIGL--------NDKMFERDFRWTDGSPLQYENWRPNQPDSFFSAGEDCVVIIWHENG-- 3439

Qy        153 AGIGGPYMFQWNDDRCNMKNNFICKYS----DEKPAVPSREAEGE 193
                 ||||  ||    : ||       : | | : :  |:
Db       3440 ---------QWNDVPCNYHLTYTCKKGTVACGQPPVVENAKTFGK 3475
```

RESULT 13
PGCA_BOVIN
ID PGCA_BOVIN STANDARD; PRT; 2364 AA.
AC P13608; P79117; Q28159;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Aggrecan core protein precursor (Cartilage-specific proteoglycan core
DE protein) (CSPCP).
GN AGC1.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RA Hering T.M., Kollar J., Huynh T.D.;
RL Submitted (SEP-1996) to the EMBL/GenBank/DDBJ databases.
RN [2]
RP SEQUENCE OF 563-1056 FROM N.A.
RX MEDLINE=89380219; PubMed=2528543;
RA Antonsson P., Heinegaard D., Oldberg A.;
RT "The keratan sulfate-enriched region of bovine cartilage proteoglycan
RT consists of a consecutively repeated hexapeptide motif.";
RL J. Biol. Chem. 264:16170-16173(1989).
RN [3]
RP SEQUENCE OF 1609-2113 AND 2151-2364 FROM N.A.
RX MEDLINE=87270630; PubMed=3111460;
RA Oldberg A., Antonsson P., Heinegaard D.;
RT "The partial amino acid sequence of bovine cartilage proteoglycan,
RT deduced from a cDNA clone, contains numerous Ser-Gly sequences
RT arranged in homologous repeats.";
RL Biochem. J. 243:255-259(1987).
RN [4]
RP SEQUENCE OF 2114-2150 FROM N.A.
RC TISSUE=Cartilage;
RX MEDLINE=93352525; PubMed=8349621;
RA Fueloep C., Walcz E., Valyon M., Glant T.T.;
RT "Expression of alternatively spliced epidermal growth factor-like
RT domains in aggrecans of different species. Evidence for a novel
RT module.";
RL J. Biol. Chem. 268:17377-17383(1993).
RN [5]
RP PARTIAL SEQUENCE.
RX MEDLINE=85027710; PubMed=6489519;
RA Perin J.-P., Bonnet F., Jolles J., Jolles P.;
RT "Sequence data concerning the protein core of the cartilage
RT proteoglycan monomers. Characterization of a sequence allowing the
RT synthesis of an oligonucleotide probe.";
RL FEBS Lett. 176:37-42(1984).
RN [6]
RP PARTIAL SEQUENCE.
RX MEDLINE=87005253; PubMed=3530809;
RA Perin J.P., Bonnet F., Jolles P.;
RT "Structural relationship between link proteins and proteoglycan
RT monomers.";
RL FEBS Lett. 206:73-77(1986).
CC -!- FUNCTION: This proteoglycan is a major component of extracellular
CC   matrix of cartilagenous tissues. A major function of this protein
CC   is to resist compression in cartilage. It binds avidly to
CC   hyaluronic acid via an amino-terminal globular region. May play a
CC   regulatory role in the matrix assembly of the cartilage.
CC -!- SUBCELLULAR LOCATION: Secreted; extracellular matrix (By
CC   similarity).
CC -!- ALTERNATIVE PRODUCTS:
CC   Event=Alternative splicing; Named isoforms=2;
CC   Name=1;
CC   IsoId=P13608-1; Sequence=Displayed;
CC   Name=2;
CC   IsoId=P13608-2; Sequence=VSP_003072;
CC -!- DOMAIN: Two globular domains, G1 and G2, comprise the amino
CC   terminus of the proteoglycan, while another globular region, G3,
CC   makes up the C-terminus. G1 contains link domains and thus
CC   consists of three disulfide-bonded loop structures designated as
CC   the A, B, B' motifs. G2 is similar to G1. The keratan sulfate (KS)
CC   and the chondroitin sulfate (CS) attachment domains lie between G2
CC   and G3.
CC -!- PTM: CONTAINS MOSTLY CHONDROITIN SULFATE, BUT ALSO N-LINKED AND O-
CC   LINKED (ABOUT 40) OLIGOSACCHARIDES.
CC -!- PTM: THE KERATAN SULFATE CONTENTS DIFFER CONSIDERABLY BETWEEN
CC   ADULT AND FETAL BOVINE PROTEOGLYCANS.
CC -!- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
CC -!- SIMILARITY: Contains 4 link domains.
CC -!- SIMILARITY: Contains 1 EGF-like domain.
CC -!- SIMILARITY: Contains 1 C-type lectin family domain.
CC -!- SIMILARITY: Contains 1 Sushi (SCR) domain.
CC -!- SIMILARITY: Belongs to the aggrecan/versican proteoglycan family.
CC ---
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC ---
DR EMBL; U76615; AAB38524.1; -.
DR EMBL; L07053; -; NOT_ANNOTATED_CDS.
DR PIR; A34234; A39808.
DR PIR; T42630; T42630.
DR HSSP; P08709; 1BF9.
DR InterPro; IPR002353; AntifreezeII.
DR InterPro; IPR000152; Asx_hydroxyl_S.
DR InterPro; IPR000742; EGF_2.
DR InterPro; IPR001881; EGF_Ca.
DR InterPro; IPR006209; EGF_like.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR001304; Lectin_C.
DR InterPro; IPR000538; Link.
DR InterPro; IPR003324; SGXXSG.
DR InterPro; IPR000436; Sushi_SCR_CCP.
DR Pfam; PF00008; EGF; 1.
DR Pfam; PF00047; ig; 1.
DR Pfam; PF00059; lectin_c; 1.
DR Pfam; PF02339; SGXXSG; 61.
DR Pfam; PF00084; sushi; 1.
DR Pfam; PF00193; Xlink; 4.
DR PRINTS; PR00356; ANTIFREEZEII.
DR PRINTS; PR01265; LINKMODULE.
DR ProDom; PD000918; Link; 4.
DR SMART; SM00032; CCP; 1.
DR SMART; SM00034; CLECT; 1.
DR SMART; SM00179; EGF_CA; 1.
DR SMART; SM00445; LINK; 4.
DR PROSITE; PS00010; ASX_HYDROXYL; 1.
DR PROSITE; PS00615; C_TYPE_LECTIN_1; 1.
DR PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS50026; EGF_3; 1.
DR PROSITE; PS01187; EGF_CA; 1.
DR PROSITE; PS50835; IG_LIKE; 1.

```
DR   PROSITE; PS00290; IG_MHC; FALSE_NEG.
DR   PROSITE; PS01241; LINK; 4.
KW   Glycoprotein; Proteoglycan; Lectin; Signal; Sushi; EGF-like domain;
KW   Calcium; Alternative splicing; Repeat; Immunoglobulin domain.
FT   SIGNAL        1     16       POTENTIAL.
FT   CHAIN        17   2364       AGGRECAN CORE PROTEIN.
FT   DOMAIN       25    147       IG-LIKE V-TYPE.
FT   DOMAIN      170    247       LINK 1.
FT   DOMAIN      268    349       LINK 2.
FT   DOMAIN      504    581       LINK 3.
FT   DOMAIN      602    683       LINK 4.
FT   DOMAIN      774    907       23 X 6 AA APPROXIMATE TANDEM REPEATS OF
FT                                E-[EK]-P-F-P-S.
FT   DOMAIN     1433   2112       CS-2.
FT   DOMAIN     2113   2149       EGF-LIKE, CALCIUM-BINDING (POTENTIAL).
FT   DOMAIN     2114   2364       G3.
FT   DOMAIN     2161   2276       C-TYPE LECTIN.
FT   DOMAIN     2280   2338       SUSHI.
FT   DISULFID     51    133       BY SIMILARITY.
FT   DISULFID    175    246       BY SIMILARITY.
FT   DISULFID    199    220       BY SIMILARITY.
FT   DISULFID    273    348       BY SIMILARITY.
FT   DISULFID    297    318       BY SIMILARITY.
FT   DISULFID    509    580       BY SIMILARITY.
FT   DISULFID    533    554       BY SIMILARITY.
FT   DISULFID    607    682       BY SIMILARITY.
FT   DISULFID    631    652       BY SIMILARITY.
FT   DISULFID   2117   2128       BY SIMILARITY.
FT   DISULFID   2182   2274       BY SIMILARITY.
FT   DISULFID   2250   2266       BY SIMILARITY.
FT   DISULFID   2281   2324       BY SIMILARITY.
FT   DISULFID   2310   2337       BY SIMILARITY.
FT   CARBOHYD    126    126       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    239    239       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    333    333       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    387    387       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    611    611       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    667    667       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   VARSPLIC   2114   2150       Missing (in isoform 2).
FT                                /FTId=VSP_003072.
SQ   SEQUENCE   2364 AA;  246359 MW;  6FF83763420C3D4C CRC64;

  Query Match             7.9%;  Score 158.5;  DB 1;  Length 2364;
  Best Local Similarity   26.8%;  Pred. No. 9.7e-05;
  Matches   48;  Conservative   29;  Mismatches   59;  Indels   43;  Gaps   10;

Qy         30 QPVCRGG---TQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFI 86
              | :|  |    |  ||:  :| |   |  : :|:  ||:    | || : :||    :|:
Db       2151 QKLCEEGWTKFQGHCYR--HFPD---RATWVDAESQCRKQQSHLSSIVTPEEQ----EFV 2201

Qy         87 ENLLPSDGDF-WIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEP----SCGSEV 141
              |   :  |: ||||         |    :  : |:||   || ||  ::|    : | :
Db       2202 NN---NAQDYQWIGL--------NDKTIEGDFRWSDGHSLQFENWRPNQPDNFFATGEDC 2250

Qy        142 CVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYS----DEKPAVPSREAEGEETE 196
               |:::|:           :|||  || :  | ||       | | |      |:: :
Db       2251 VVMIWHEKG----------EWNDVPCNYQLPFTCKKGTVACGEPPVVEHARIFGQKKD 2298


RESULT 14
PGCA_HUMAN
ID   PGCA_HUMAN     STANDARD;      PRT;  2415 AA.
AC   P16112; Q13650;
DT   01-APR-1990 (Rel. 14, Created)
DT   01-AUG-1992 (Rel. 23, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   Aggrecan core protein precursor (Cartilage-specific proteoglycan core
DE   protein) (CSPCP) (Chondroitin sulfate proteoglycan core protein 1).
GN   AGC1 OR CSPG1.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORM 3).
RC   TISSUE=Chondrocytes;
RX   MEDLINE=91093289; PubMed=1985970;
RA   Doege K.J., Sasaki M., Kimura T., Yamada Y.;
RT   "Complete coding sequence and deduced primary structure of the human
RT   cartilage large aggregating proteoglycan, aggrecan. Human-specific
RT   repeats, and additional alternatively spliced forms.";
RL   J. Biol. Chem. 266:894-902(1991).
RN   [2]
RP   SEQUENCE OF 1778-2415 FROM N.A. (ISOFORM 2).
RC   TISSUE=Chondrocytes;
RA   Dudhia J., Hardingham T.E.;
RL   Submitted (JAN-1990) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE OF 1936-2415 FROM N.A. (ISOFORM 1).
RX   MEDLINE=89380154; PubMed=2789216;
RA   Baldwin C.T., Reginato A.M., Prockop D.J.;
RT   "A new epidermal growth factor-like domain in the human core protein
RT   for the large cartilage-specific proteoglycan. Evidence for
RT   alternative splicing of the domain.";
RL   J. Biol. Chem. 264:15747-15750(1989).
RN   [4]
RP   SEQUENCE OF 764-864 FROM N.A.
RC   TISSUE=Blood;
RX   MEDLINE=95128522; PubMed=7827755;
RA   Barry F.P., Neame P.J., Sasse J., Pearson D.;
RT   "Length variation in the keratan sulfate domain of mammalian
RT   aggrecan.";
RL   Matrix Biol. 14:323-328(1994).
CC   -!- FUNCTION: THIS PROTEOGLYCAN IS A MAJOR COMPONENT OF EXTRACELLULAR
CC       MATRIX OF CARTILAGENOUS TISSUES. A MAJOR FUNCTION OF THIS PROTEIN
CC       IS TO RESIST COMPRESSION IN CARTILAGE. IT BINDS AVIDLY TO
CC       HYALURONIC ACID VIA AN AMINO-TERMINAL GLOBULAR REGION.
CC   -!- SUBUNIT: Interacts with FBLN1 (By similarity).
CC   -!- SUBCELLULAR LOCATION: Secreted; extracellular matrix (By
CC       similarity).
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=3;
CC         Comment=Additional isoforms seem to exist;
CC       Name=1;
CC         IsoId=P16112-1; Sequence=Displayed;
CC       Name=2;
CC         IsoId=P16112-2; Sequence=VSP_003074;
CC       Name=3;
CC         IsoId=P16112-3; Sequence=VSP_003074, VSP_003075;
CC   -!- DOMAIN: Two globular domains, G1 and G2, comprise the amino
CC       terminus of the proteoglycan, while another globular region, G3,
CC       makes up the COOH terminus. G1 contains link domains and thus
CC       consists of three disulfide-bonded loop structures designated as
CC       the A, B, B' motifs. G2 is similar to G1. The keratan sulfate (KS)
CC       and the chondroitin sulfate (CS) attachment domains lie between G2
CC       and G3.
CC   -!- PTM: Contains mostly chondroitin sulfate, but also keratan sulfate
CC       chains, N-linked and O-linked oligosaccharides.
CC   -!- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
CC   -!- SIMILARITY: Contains 4 link domains.
CC   -!- SIMILARITY: Contains 1 EGF-like domain.
CC   -!- SIMILARITY: Contains 1 C-type lectin family domain.
CC   -!- SIMILARITY: Contains 1 Sushi (SCR) domain.
CC   -!- SIMILARITY: Belongs to the aggrecan/versican proteoglycan family.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; M55172; AAA62824.1; -.
DR   EMBL; J05062; AAA35726.1; -.
DR   EMBL; X17406; CAA35463.1; -.
```

```
DR   EMBL; S74659; AAC60643.2; -.
DR   PIR; A39086; A39086.
DR   HSSP; P98066; 1TSG.
DR   Genew; HGNC:319; AGC1.
DR   MIM; 155760; -.
DR   InterPro; IPR006209; EGF_like.
DR   InterPro; IPR006210; IEGF.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003599; Ig.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR001304; Lectin_C.
DR   InterPro; IPR000538; Link.
DR   InterPro; IPR003324; SGXXSG.
DR   InterPro; IPR000436; Sushi_SCR_CCP.
DR   Pfam; PF00008; EGF; 1.
DR   Pfam; PF00047; ig; 1.
DR   Pfam; PF00059; lectin_c; 1.
DR   Pfam; PF02339; SGXXSG; 71.
DR   Pfam; PF00084; sushi; 1.
DR   Pfam; PF00193; Xlink; 4.
DR   PRINTS; PR01265; LINKMODULE.
DR   ProDom; PD000918; Link; 4.
DR   SMART; SM00032; CCP; 1.
DR   SMART; SM00034; CLECT; 1.
DR   SMART; SM00181; EGF; 1.
DR   SMART; SM00409; IG; 1.
DR   SMART; SM00445; LINK; 4.
DR   PROSITE; PS00615; C_TYPE_LECTIN_1; 1.
DR   PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
DR   PROSITE; PS00022; EGF_1; 1.
DR   PROSITE; PS01186; EGF_2; 1.
DR   PROSITE; PS50026; EGF_3; 1.
DR   PROSITE; PS50835; IG_LIKE; 1.
DR   PROSITE; PS00290; IG_MHC; 1.
DR   PROSITE; PS01241; LINK; 3.
KW   Glycoprotein; Proteoglycan; Lectin; Signal; Sushi; EGF-like domain;
KW   Alternative splicing; Repeat; Immunoglobulin domain.
FT   SIGNAL        1     19       POTENTIAL.
FT   CHAIN        20   2415       AGGRECAN CORE PROTEIN.
FT   DOMAIN       34    147       IG-LIKE V-TYPE.
FT   DOMAIN      170    247       LINK 1.
FT   DOMAIN      268    349       LINK 2.
FT   DOMAIN      495    572       LINK 3.
FT   DOMAIN      593    673       LINK 4.
FT   DOMAIN     2164   2199       EGF-LIKE.
FT   DOMAIN     2201   2327       C-TYPE LECTIN.
FT   DOMAIN     2331   2389       SUSHI.
FT   DOMAIN       48    141       G1-A.
FT   DOMAIN      152    247       G1-B.
FT   DOMAIN      253    349       G1-B'.
FT   DOMAIN      477    571       G2-B.
FT   DOMAIN      578    672       G2-B'.
FT   DOMAIN      676    848       KS.
FT   DOMAIN      772    844       12 X APPROXIMATE TANDEM REPEATS.
FT   DOMAIN      851   1497       CS-1.
FT   DOMAIN      941   1497       29 X APPROXIMATE TANDEM REPEATS.
FT   DOMAIN     1498   2162       CS-2.
FT   DOMAIN     2163   2415       G3.
FT   DISULFID     51    133       BY SIMILARITY.
FT   DISULFID    175    246       BY SIMILARITY.
FT   DISULFID    199    220       BY SIMILARITY.
FT   DISULFID    273    348       BY SIMILARITY.
FT   DISULFID    297    318       BY SIMILARITY.
FT   DISULFID    500    571       BY SIMILARITY.
FT   DISULFID    524    545       BY SIMILARITY.
FT   DISULFID    598    672       BY SIMILARITY.
FT   DISULFID    621    642       BY SIMILARITY.
FT   DISULFID   2168   2178       BY SIMILARITY.
FT   DISULFID   2173   2187       BY SIMILARITY.
FT   DISULFID   2189   2198       BY SIMILARITY.
FT   DISULFID   2205   2216       BY SIMILARITY.
FT   DISULFID   2233   2325       BY SIMILARITY.
FT   DISULFID   2301   2317       BY SIMILARITY.
FT   DISULFID   2332   2375       BY SIMILARITY.
FT   DISULFID   2361   2388       BY SIMILARITY.
FT   CARBOHYD    126    126       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    239    239       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    333    333       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    387    387       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    434    434       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    602    602       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    657    657       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    737    737       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1898   1898       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   VARSPLIC   2163   2200       Missing (in isoform 2 and isoform 3).
FT                                /FTId=VSP_003074.
FT   VARSPLIC   2330   2390       Missing (in isoform 3).
FT                                /FTId=VSP_003075.
FT   CONFLICT    766    766       E -> A (IN REF. 4).
FT   CONFLICT    847    847       E -> V (IN REF. 4).
FT   CONFLICT   1928   1928       E -> A (IN REF. 2).
FT   CONFLICT   1964   1964       I -> V (IN REF. 2 AND 3).
FT   CONFLICT   2070   2070       P -> A (IN REF. 3).
FT   CONFLICT   2391   2391       A -> P (IN REF. 2 AND 3).
SQ   SEQUENCE   2415 AA;  250191 MW;  1288937E1B98C6B6 CRC64;
```

Query Match 7.9%; Score 158.5; DB 1; Length 2415;
Best Local Similarity 29.1%; Pred. No. 9.9e-05;
Matches 52; Conservative 22; Mismatches 62; Indels 43; Gaps 11;

```
Qy        30 QPVCRGG---TQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFI 86
             | ||  |    |  ||:  :| |   |  : :|:  ||     | || : :||    :|:
Db      2202 QEVCEEGWNKYQGHCYR--HFPD---RETWVDAERRCREQQSHLSSIVTPEEQ----EFV 2252

Qy        87 ENLLPSDGDF-WIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEPS---CGSEVC 142
               |   :  |: ||||         |    :  : |:||   || ||  ::|       | |
Db      2253 NN---NAQDYQWIGL--------NDRTIEGDFRWSDGHPMQFENWRPNQPDNFFAAGEDC 2301

Qy       143 VVM-YHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYS----DEKPAVPSREAEGEETE 196
             ||| :|:             :||| ||   | ||       | | |      |:: :
Db      2302 VVMIWHEKG----------EWNDVPCNYHLPFTCKKGTVACGEPPVVEHARTFGQKKD 2349
```

```
RESULT 15
CD93_HUMAN
ID   CD93_HUMAN     STANDARD;      PRT;   652 AA.
AC   Q9NPY3; O00274;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Complement component C1q receptor precursor (Complement component 1, q
DE   subcomponent, receptor 1) (C1qRp) (C1qR(p)) (C1q/MBL/SPA receptor)
DE   (CD93 antigen) (CDw93).
GN   C1QR1 OR CD93.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX   MEDLINE=97199258; PubMed=9047234;
RA   Nepomuceno R.R., Henschen-Edman A.H., Burgess W.H., Tenner A.J.;
RT   "cDNA cloning and primary structure analysis of C1qR(P), the human
RT   C1q/MBL/SPA receptor that mediates enhanced phagocytosis in vitro.";
RL   Immunity 6:119-129(1997).
RN   [2]
RP   SEQUENCE FROM N.A., AND VARIANT ALA-318.
RX   MEDLINE=21640567; PubMed=11781389;
RA   Steinberger P., Szekeres A., Wille S., Stockl J., Selenko N.,
RA   Prager E., Staffler G., Madic O., Stockinger H., Knapp W.;
RT   "Identification of human CD93 as the phagocytic C1q receptor (C1qRp)
RT   by expression cloning.";
RL   J. Leukoc. Biol. 71:133-140(2002).
RN   [3]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21638749; PubMed=11780052;
```

```
RA   Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,
RA   Jones M., Stavrides G., Almeida J.P., Babbage A.K., Bagguley C.L.,
RA   Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,
RA   Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,
RA   Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,
RA   Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,
RA   Clegg S., Cobley V.E., Collier R.E., Connor R.E., Corby N.R.,
RA   Coulson A., Coville G.J., Deadman R., Dhami P.D., Dunn M.,
RA   Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,
RA   Grafham D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,
RA   Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,
RA   Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,
RA   Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,
RA   Lehvaeslaiho M.H., Leversha M.A., Lloyd C., Lloyd D.M., Lovell J.D.,
RA   Marsh V.L., Martin S.L., McConnachie L.J., McLay K., McMurray A.A.,
RA   Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,
RA   Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,
RA   Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsay H.,
RA   Rice C.M., Ross M.T., Scott C.E., Sehra H.K., Shownkeen R., Sims S.,
RA   Skuce C.D., Smith M.L., Soderlund C., Steward C.A., Sulston J.E.,
RA   Swann R.M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,
RA   Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M.,
RA   Whitehead S.L., Whittaker P., Willey D.L., Williams L., Williams S.A.,
RA   Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,
RA   Rogers J.;
RT   "The DNA sequence and comparative analysis of human chromosome 20.";
RL   Nature 414:865-871(2001).
RN   [4]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Leukocyte;
RX   MEDLINE=22388257; PubMed=12477932;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length
RT   human and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [5]
RP   CHARACTERIZATION.
RX   MEDLINE=21990337; PubMed=11994479;
RA   McGreal E.P., Ikewaki N., Akatsu H., Morgan B.P., Gasque P.;
RT   "Human C1qRp is identical with CD93 and the mNI-11 antigen but does
RT   not bind C1q.";
RL   J. Immunol. 168:5222-5232(2002).
RN   [6]
RP   O-GLYCOSYLATION.
RX   MEDLINE=99192777; PubMed=10092817;
RA   Nepomuceno R.R., Ruiz S., Park M., Tenner A.J.;
RT   "C1qRP is a heavily O-glycosylated cell surface protein involved in
RT   the regulation of phagocytic activity.";
RL   J. Immunol. 162:3583-3589(1999).
CC   -!- FUNCTION: Receptor (or element of a larger receptor complex) for
CC       C1q, mannose-binding lectin (MBL2) and pulmonary surfactant
CC       protein A (SPA). May mediate the enhancement of phagocytosis in
CC       monocytes and macrophages upon interaction with soluble defense
CC       collagens. May play a role in intercellular adhesion.
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC   -!- TISSUE SPECIFICITY: Highly expressed in endothelial cells,
CC       platelets, cells of myeloid origin, such as monocytes and
CC       neutrophils. Not expressed in cells of lymphoid origin.
CC   -!- PTM: N- and O-glycosylated.
CC   -!- SIMILARITY: Contains 1 C-type lectin family domain.
CC   -!- SIMILARITY: Contains 5 EGF-like domains.
CC   -!- CAUTION: Has been sometimes referred to as a collectin receptor.
CC   -!- CAUTION: According to Ref.5, C1q is not a ligand for C1QR1.
CC   -!- DATABASE: NAME=PROW; NOTE=PROW 3:1-6(2001);
CC       WWW="http://www.ncbi.nlm.nih.gov/prow/guide/467246456_g.htm".
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U94333; AAB53110.1; -.
DR   EMBL; AL118508; CAC00597.1; -.
DR   EMBL; BC028075; AAH28075.1; -.
DR   HSSP; P35555; 1EMN.
DR   Genew; HGNC:15855; C1QR1.
DR   MIM; 120577; -.
DR   GO; GO:0016021; C:integral to membrane; IC.
DR   GO; GO:0004872; F:receptor activity; NAS.
DR   GO; GO:0016337; P:cell-cell adhesion; IDA.
DR   GO; GO:0042116; P:macrophage activation; NAS.
DR   GO; GO:0006909; P:phagocytosis; NAS.
DR   InterPro; IPR000152; Asx_hydroxyl_S.
DR   InterPro; IPR001881; EGF_Ca.
DR   InterPro; IPR006209; EGF_like.
DR   InterPro; IPR001304; Lectin_C.
DR   Pfam; PF00008; EGF; 5.
DR   Pfam; PF00059; lectin_c; 1.
DR   SMART; SM00034; CLECT; 1.
DR   SMART; SM00179; EGF_CA; 3.
DR   PROSITE; PS00010; ASX_HYDROXYL; 3.
DR   PROSITE; PS00615; C_TYPE_LECTIN_1; FALSE_NEG.
DR   PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
DR   PROSITE; PS01186; EGF_2; 3.
DR   PROSITE; PS50026; EGF_3; 3.
DR   PROSITE; PS01187; EGF_CA; 3.
KW   Cell adhesion; Receptor; Repeat; Signal; Transmembrane;
KW   EGF-like domain; Lectin; Glycoprotein; Polymorphism.
FT   SIGNAL        1     21
FT   CHAIN        22    652       COMPLEMENT COMPONENT C1Q RECEPTOR.
FT   DOMAIN       24    580       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    581    601       POTENTIAL.
FT   DOMAIN      602    652       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN       32    174       C-TYPE LECTIN.
FT   DOMAIN      260    301       EGF-LIKE 1.
FT   DOMAIN      302    344       EGF-LIKE 2.
FT   DOMAIN      345    384       EGF-LIKE 3, CALCIUM-BINDING (POTENTIAL).
FT   DOMAIN      385    426       EGF-LIKE 4, CALCIUM-BINDING (POTENTIAL).
FT   DOMAIN      427    468       EGF-LIKE 5, CALCIUM-BINDING (POTENTIAL).
FT   DOMAIN      594    601       POLY-LEU.
FT   DISULFID    264    275       BY SIMILARITY.
FT   DISULFID    271    285       BY SIMILARITY.
FT   DISULFID    287    300       BY SIMILARITY.
FT   DISULFID    306    317       BY SIMILARITY.
FT   DISULFID    311    328       BY SIMILARITY.
FT   DISULFID    330    343       BY SIMILARITY.
FT   DISULFID    349    358       BY SIMILARITY.
FT   DISULFID    354    367       BY SIMILARITY.
FT   DISULFID    369    383       BY SIMILARITY.
FT   DISULFID    389    400       BY SIMILARITY.
FT   DISULFID    396    409       BY SIMILARITY.
FT   DISULFID    411    425       BY SIMILARITY.
FT   DISULFID    431    443       BY SIMILARITY.
FT   DISULFID    439    452       BY SIMILARITY.
FT   DISULFID    454    467       BY SIMILARITY.
FT   CARBOHYD    325    325       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   VARIANT     318    318       V -> A.
```

```
FT                                /FTId=VAR_013573.
FT   CONFLICT     22     22       T -> V (IN REF. 1; AA SEQUENCE).
FT   CONFLICT     36     36       C -> T (IN REF. 1; AA SEQUENCE).
FT   CONFLICT     38     39       TA -> RI (IN REF. 1; AA SEQUENCE).
FT   CONFLICT    155    155       S -> N (IN REF. 1).
FT   CONFLICT    186    186       G -> A (IN REF. 1; AA SEQUENCE).
FT   CONFLICT    492    492       S -> A (IN REF. 1; AA SEQUENCE).
FT   CONFLICT    496    496       R -> Q (IN REF. 1; AA SEQUENCE).
FT   CONFLICT    504    504       R -> G (IN REF. 1; AA SEQUENCE).
FT   CONFLICT    541    541       P -> S (IN REF. 1).
SQ   SEQUENCE   652 AA;  68560 MW;  EECA0FEAC55FCAC2 CRC64;

  Query Match             7.9%;  Score 157.5;  DB 1;  Length 652;
  Best Local Similarity   25.2%;  Pred. No. 2e-05;
  Matches   53;  Conservative   40;  Mismatches   84;  Indels   33;  Gaps    9;

Qy         10 VLLAVLLVGLRAATGRLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQ 69
              :|| :||    | ||       || |     ||        | :|:  ||:  | ::||
Db          9 LLLLLLLTQPGAGTGADTEA-VVCVG---TACYTA-----HSGKLSAAEAQNHCNQNGGN 59

Qy         70 LVSIESEDEQKLIEKFIENLLPSD-------GDFWIGLRRREEKQSNSTACQDLYAWT-D 121
              | :::|::| : ::: :  ||  :        |||||:| : |  : :     ::|
Db         60 LATVKSKEEAQHVQRVLAQLLRREAALTARMSKFWIGLQREKGKCLDPSLPLKGFSWVGG 119

Qy        122 GSISQFRNWYVD-EPSCGSEVCVVM---YHQPSAPAGIGGPYMFQWNDDRCNMKNN---- 173
              |  : : ||: :   || |: ||  :     ||  |: :      :|::  |    :
Db        120 GEDTPYSNWHKELRNSCISKRCVSLLLDLSQPLLPSRLP-----KWSEGPCGSPGSPGSN 174

Qy        174 ---FICKYSDEKPAVPSREAEGEETELTTP 200
                 |:||:| :    |       :   |||
Db        175 IEGFVCKFSFKGMCRPLALGGPGQVTYTTP 204


Search completed: September  9, 2004, 22:21:45
Job time : 14 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: September 9, 2004, 22:19:06 ; Search time 43 Seconds
(without alignments)
2744.275 Million cell updates/sec

Title: US-09-887-855-2
Perfect score: 2000
Sequence: 1 MRPGTALQAVLLAVLLVGLR.....PDQMGRSKESGWVENEIYGY 374

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_25:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2000	100.0	374	4	Q96NF3	Q96nf3 homo sapien
2	1997	99.9	374	4	Q8TAY8	Q8tay8 homo sapien
3	1996	99.8	374	4	Q96NC5	Q96nc5 homo sapien
4	1697.5	84.9	374	11	Q9Z209	Q9z209 cricetulus
5	862.5	43.1	211	11	Q8C351	Q8c351 mus musculu
6	652	32.6	292	11	Q8BVU2	Q8bvu2 mus musculu
7	646	32.3	246	11	Q8BMI7	Q8bmi7 mus musculu
8	576	28.8	232	4	Q7Z799	Q7z799 homo sapien
9	513	25.7	236	4	Q7Z798	Q7z798 homo sapien
10	500	25.0	236	4	Q7Z7A0	Q7z7a0 homo sapien
11	185	9.2	1290	13	Q9W6E1	Q9w6e1 gallus gall
12	178.5	8.9	1456	11	Q61830	Q61830 mus musculu
13	177.5	8.9	1348	5	Q25199	Q25199 hydra atten
14	176.5	8.8	1479	4	Q9Y5P9	Q9y5p9 homo sapien
15	176.5	8.8	1479	4	Q9UBG0	Q9ubg0 homo sapien
16	170.5	8.5	1479	11	Q64449	Q64449 mus musculu
17	169	8.5	217	11	Q8C4F8	Q8c4f8 mus musculu
18	165.5	8.3	742	11	Q8K4Q8	Q8k4q8 mus musculu
19	165.5	8.3	742	11	Q8C979	Q8c979 mus musculu
20	161.5	8.1	134	5	Q9XYX3	Q9xyx3 hydra magni
21	159.5	8.0	742	11	Q8VIF6	Q8vif6 mus musculu
22	159	8.0	1031	5	Q8WSX2	Q8wsx2 dugesia tig
23	158	7.9	142	11	Q8CJ86	Q8cj86 mus musculu
24	158	7.9	142	11	Q8BHK7	Q8bhk7 mus musculu
25	158	7.9	295	11	Q91ZW4	Q91zw4 mus musculu
26	158	7.9	311	11	Q9D8V4	Q9d8v4 mus musculu
27	158	7.9	325	11	Q91ZX0	Q91zx0 mus musculu
28	157.5	7.9	358	6	Q8HY04	Q8hy04 papio hamad
29	157.5	7.9	381	6	Q8SQB2	Q8sqb2 macaca mula
30	157.5	7.9	652	4	Q8IXK1	Q8ixk1 homo sapien
31	157	7.8	158	13	Q90WI7	Q90wi7 bungarus fa
32	156	7.8	162	5	Q25459	Q25459 megabalanus
33	156	7.8	323	11	Q8CJ91	Q8cj91 mus musculu
34	156	7.8	339	6	Q95244	Q95244 sus scrofa
35	155	7.8	293	11	Q8BGZ0	Q8bgz0 mus musculu
36	155	7.8	323	11	Q8CJ94	Q8cj94 mus musculu
37	155	7.8	323	11	Q8CJ93	Q8cj93 mus musculu
38	155	7.8	323	11	Q8CJ88	Q8cj88 mus musculu
39	155	7.8	379	11	Q7TMA7	Q7tma7 mus musculu
40	155	7.8	381	6	Q95LA8	Q95la8 macaca mula
41	155	7.8	473	11	Q7TSP9	Q7tsp9 mus musculu
42	155	7.8	477	11	Q7TSQ7	Q7tsq7 mus musculu
43	155	7.8	504	11	Q7TSQ0	Q7tsq0 mus musculu
44	155	7.8	534	11	Q7TSQ1	Q7tsq1 mus musculu
45	155	7.8	1152	13	Q90WM2	Q90wm2 xenopus lae

ALIGNMENTS

RESULT 1
Q96NF3
ID Q96NF3 PRELIMINARY; PRT; 374 AA.
AC Q96NF3;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein FLJ30977.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Ishibashi T., Kanehori K., Yosida M., Watanabe S., Ishida S., Ono Y.,
RA Hotuta T., Hiraoka S., Murakawa K., Takiguchi S., Kusano J.,
RA Watanabe M., Fujimori K., Tanai H., Ishida M., Yamashita H., Chiba Y.,
RA Sugiyama T., Irie R., Otsuki T., Sato H., Wakamatsu A., Ishii S.,
RA Yamamoto J., Isono Y., Kawai-Hio Y., Saito K., Nishikawa T.,
RA Kimura K., Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K.,
RA Wagatsuma M., Takahashi-Fujii A., Oshima A., Sugiyama A., Kawakami B.,
RA Suzuki Y., Sugano S., Nagahari K., Masuho Y., Nagai K., Isogai T.;
RT "NEDO human cDNA sequencing project.";
RL Submitted (OCT-2001) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AK055539; BAB70946.1; -.
DR GO; GO:0005529; F:sugar binding; IEA.
DR InterPro; IPR001304; Lectin_C.
DR Pfam; PF00059; lectin_c; 1.
DR SMART; SM00034; CLECT; 1.
DR PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
KW Hypothetical protein.
SQ SEQUENCE 374 AA; 42280 MW; 8AE64E6BC9E56DCD CRC64;

Query Match 100.0%; Score 2000; DB 4; Length 374;
Best Local Similarity 100.0%; Pred. No. 3e-179;
Matches 374; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MRPGTALQAVLLAVLLVGLRAATGRLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAK 60

```
Db          1 MRPGTALQAVLLAVLLVGLRAATGRLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAK 60

Qy         61 EACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 EACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWT 120

Qy        121 DGSISQFRNWYVDEPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSD 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 DGSISQFRNWYVDEPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSD 180

Qy        181 EKPAVPSREAEGEETELTTPVLPEETQEEDAKKTFKESREAALNLAYILIPSIPLLLLLV 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 EKPAVPSREAEGEETELTTPVLPEETQEEDAKKTFKESREAALNLAYILIPSIPLLLLLV 240

Qy        241 VTTVVCWVWICRKRKREQPDPSTKKQHTIWPSPHQGNSPDLEVYNVIRKQSEADLAETRP 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 VTTVVCWVWICRKRKREQPDPSTKKQHTIWPSPHQGNSPDLEVYNVIRKQSEADLAETRP 300

Qy        301 DLKNISFRVCSGEATPDDMSCDYDNMAVNPSESGFVTLVSVESGFVTNDIYEFSPDQMGR 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 DLKNISFRVCSGEATPDDMSCDYDNMAVNPSESGFVTLVSVESGFVTNDIYEFSPDQMGR 360

Qy        361 SKESGWVENEIYGY 374
              ||||||||||||||
Db        361 SKESGWVENEIYGY 374
```

RESULT 2
Q8TAY8
ID Q8TAY8 PRELIMINARY; PRT; 374 AA.
AC Q8TAY8;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Similar to unnamed protein product.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Strausberg R.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DDBJ databases.
DR EMBL; BC025407; AAH25407.1; -.
DR GO; GO:0005529; F:sugar binding; IEA.
DR InterPro; IPR001304; Lectin_C.
DR Pfam; PF00059; lectin_c; 1.
DR SMART; SM00034; CLECT; 1.
DR PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
SQ SEQUENCE 374 AA; 42312 MW; FC214E6BC9E578D9 CRC64;

Query Match 99.9%; Score 1997; DB 4; Length 374;
Best Local Similarity 99.7%; Pred. No. 5.7e-179;
Matches 373; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MRPGTALQAVLLAVLLVGLRAATGRLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAK 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MRPGTALQAVLLAVLLVGLRAATGRLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAK 60

Qy         61 EACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 EACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWT 120

Qy        121 DGSISQFRNWYVDEPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSD 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 DGSISQFRNWYVDEPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSD 180

Qy        181 EKPAVPSREAEGEETELTTPVLPEETQEEDAKKTFKESREAALNLAYILIPSIPLLLLLV 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 EKPAVPSREAEGEETELTTPVLPEETQEEDAKKTFKESREAALNLAYILIPSIPLLLLLV 240

Qy        241 VTTVVCWVWICRKRKREQPDPSTKKQHTIWPSPHQGNSPDLEVYNVIRKQSEADLAETRP 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 VTTVVCWVWICRKRKREQPDPSTKKQHTIWPSPHQGNSPDLEVYNVIRKQSEADLAETRP 300

Qy        301 DLKNISFRVCSGEATPDDMSCDYDNMAVNPSESGFVTLVSVESGFVTNDIYEFSPDQMGR 360
              ||||||||||||||||||||||||||||||||||||||||||||:||||||||||||||||
Db        301 DLKNISFRVCSGEATPDDMSCDYDNMAVNPSESGFMTLVSVESGFVTNDIYEFSPDQMGR 360

Qy        361 SKESGWVENEIYGY 374
              ||||||||||||||
Db        361 SKESGWVENEIYGY 374
```

RESULT 3
Q96NC5
ID Q96NC5 PRELIMINARY; PRT; 374 AA.
AC Q96NC5;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein FLJ31092.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Tashiro H., Yamazaki M., Watanabe K., Kumagai A., Itakura S.,
RA Fukuzumi Y., Fujimori Y., Komiyama M., Sugiyama T., Irie R.,
RA Otsuki T., Sato H., Wakamatsu A., Ishii S., Yamamoto J., Isono Y.,
RA Kawai-Hio Y., Saito K., Nishikawa T., Kimura K., Yamashita H.,
RA Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K., Wagatsuma M.,
RA Murakawa K., Kanehori K., Takahashi-Fujii A., Oshima A., Sugiyama A.,
RA Kawakami B., Suzuki Y., Sugano S., Nagahari K., Masuho Y., Nagai K.,
RA Isogai T.;
RT "NEDO human cDNA sequencing project.";
RL Submitted (OCT-2001) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AK055654; BAB70978.1; -.
DR GO; GO:0005529; F:sugar binding; IEA.
DR InterPro; IPR001304; Lectin_C.
DR Pfam; PF00059; lectin_c; 1.
DR SMART; SM00034; CLECT; 1.
DR PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
KW Hypothetical protein.
SQ SEQUENCE 374 AA; 42310 MW; CBF74E676E23BA16 CRC64;

Query Match 99.8%; Score 1996; DB 4; Length 374;
Best Local Similarity 99.7%; Pred. No. 7.1e-179;
Matches 373; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 MRPGTALQAVLLAVLLVGLRAATGRLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAK 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MRPGTALQAVLLAVLLVGLRAATGRLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAK 60

Qy         61 EACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 EACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWT 120

Qy        121 DGSISQFRNWYVDEPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSD 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 DGSISQFRNWYVDEPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSD 180

Qy        181 EKPAVPSREAEGEETELTTPVLPEETQEEDAKKTFKESREAALNLAYILIPSIPLLLLLV 240
              |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db        181 EKPAVPSREAEGEETELTTPVLPEETQEEDTKKTFKESREAALNLAYILIPSIPLLLLLV 240

Qy        241 VTTVVCWVWICRKRKREQPDPSTKKQHTIWPSPHQGNSPDLEVYNVIRKQSEADLAETRP 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 VTTVVCWVWICRKRKREQPDPSTKKQHTIWPSPHQGNSPDLEVYNVIRKQSEADLAETRP 300

Qy        301 DLKNISFRVCSGEATPDDMSCDYDNMAVNPSESGFVTLVSVESGFVTNDIYEFSPDQMGR 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 DLKNISFRVCSGEATPDDMSCDYDNMAVNPSESGFVTLVSVESGFVTNDIYEFSPDQMGR 360
```

```
Qy        361 SKESGWVENEIYGY 374
              ||||||||||||||
Db        361 SKESGWVENEIYGY 374


RESULT 4
Q9Z209
ID   Q9Z209      PRELIMINARY;      PRT;   374 AA.
AC   Q9Z209;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Layilin.
OS   Cricetulus griseus (Chinese hamster).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC   Cricetulus.
OX   NCBI_TaxID=10029;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Borowsky M.L., Hynes R.O.;
RT   "Layilin, a novel talin-binding transmembrane protein homologous with
RT   C-type lectins, is localized in membrane ruffles.";
RL   J. Cell Biol. 143:0-0(1998).
DR   EMBL; AF093673; AAC68695.1; -.
DR   HSSP; P22897; 1EGG.
DR   GO; GO:0005529; F:sugar binding; IEA.
DR   InterPro; IPR001304; Lectin_C.
DR   Pfam; PF00059; lectin_c; 1.
DR   SMART; SM00034; CLECT; 1.
DR   PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
SQ   SEQUENCE   374 AA;  42435 MW;  298A8BA24FB04E1C CRC64;

  Query Match             84.9%;  Score 1697.5;  DB 11;  Length 374;
  Best Local Similarity   84.7%;  Pred. No. 7.8e-151;
  Matches  316;  Conservative   23;  Mismatches   33;  Indels    1;  Gaps    1;

Qy          1 MRPGTALQAVLLAVLLVGLRAATGRLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAK 60
              |:|| |||||||||||   |:: ||||||| ||||||:|||||||||||  :||||||||
Db          1 MQPGPALQAVLLAVLLSEPRSSKGRLLSGQLVCRGGTRRPCYKVIYFHDAFQRLNFEEAK 60

Qy         61 EACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWT 120
              ||||||||||||||:||||:|||||||||| |||||||||| | || |:||||||||||
Db         61 EACRRDGGQLVSIETEDEQRLIEKFIENLLASDGDFWIGLRRLEVKQVNNTACQDLYAWT 120

Qy        121 DGSISQFRNWYVDEPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSD 180
              ||| |||||||||||||||||||||||||| |||| |||||||||||||||||||:|
Db        121 DGSTSQFRNWYVDEPSCGSEVCVVMYHQPSAPPGIGGSYMFQWNDDRCNMKNNFICKYAD 180

Qy        181 EKPA-VPSREAEGEETELTTPVLPEETQEEDAKKTFKESREAALNLAYILIPSIPLLLLL 239
              |||:  ||    || ||  |||||||||:|| |:|||||||||||||||||||| |||
Db        181 EKPSTTPSIRPGGEATEPPTPVLPEETQKEDTKETFKESREAALNLAYILIPSIPLFLLL 240

Qy        240 VVTTVVCWVWICRKRKREQPDPSTKKQHTIWPSPHQGNSPDLEVYNVIRKQSEADLAETR 299
              |||:  |||||||:||:|||||:||:||||||:||| |||:|:|||||||||||| | |
Db        241 VVTSAACWVWICRRRKQEQPDPTTKEQHTIWPTPHQENSPNLDVYNVIRKQSEADLTEPR 300

Qy        300 PDLKNISFRVCSGEATPDDMSCDYDNMAVNPSESGFVTLVSVESGFVTNDIYEFSPDQMG 359
              |||||||||||| || |||:|||||||||||||||| |:|||||||||||||:||
Db        301 PDLKNISFRVCSSEAPPDDISCDYDNMAVNPSESGFVTLASMESGFVTNDIYEFSPDRMG 360

Qy        360 RSKESGWVENEIY 372
              |||||||||||||
Db        361 RSKESGWVENEIY 373


RESULT 5
Q8C351
ID   Q8C351      PRELIMINARY;      PRT;   211 AA.
AC   Q8C351;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   LAYILIN homolog (Fragment).
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Lung;
RX   MEDLINE=22354683; PubMed=12466851;
RA   The FANTOM Consortium,
RA   the RIKEN Genome Exploration Research Group Phase I & II Team;
RT   "Analysis of the mouse transcriptome based on functional annotation of
RT   60,770 full-length cDNAs.";
RL   Nature 420:563-573(2002).
DR   EMBL; AK086930; BAC39765.1; -.
DR   GO; GO:0005529; F:sugar binding; IEA.
DR   InterPro; IPR001304; Lectin_C.
DR   Pfam; PF00059; lectin_c; 1.
DR   SMART; SM00034; CLECT; 1.
DR   PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
FT   NON_TER     211     211
SQ   SEQUENCE   211 AA;  23697 MW;  AD9870B5957DD5AE CRC64;

  Query Match             43.1%;  Score 862.5;  DB 11;  Length 211;
  Best Local Similarity   77.7%;  Pred. No. 1e-72;
  Matches  164;  Conservative   12;  Mismatches   26;  Indels    9;  Gaps    2;

Qy          1 MRPGTALQAVLLAVLLVGLRAATGRLLS--------GQPVCRGGTQRPCYKVIYFHDTSR 52
              |:|| ||||:||||||    | : |||||         || |||||||||||||||  :
Db          1 MQPGAALQAMLLAVLLAKPRDSKGRLLSASDLDPRGGQLVCRGGTRRPCYKVIYFHDAFQ 60

Qy         53 RLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTA 112
              ||||||||| |  ||||||||||:||||:|||||||||| |||||||||:| |||||||||
Db         61 RLNFEEAKETCMEDGGQLVSIETEDEQRLIEKFIENLLASDGDFWIGLKRLEEKQSNNTA 120

Qy        113 CQDLYAWTDGSISQFRNWYVDEPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKN 172
              |||||||||| |||||||||||||||||||||||||||||||| |||| ||||||||||||
Db        121 CQDLYAWTDGSTSQFRNWYVDEPSCGSEVCVVMYHQPSAPPGIGGSYMFQWNDDRCNMKN 180

Qy        173 NFICKYSDEKPA-VPSREAEGEETELTTPVL 202
              |||||| |:||:  ||    || ||  ||:|
Db        181 NFICKYHDDKPSTTPSIWPGGEATEPATPLL 211


RESULT 6
Q8BVU2
ID   Q8BVU2      PRELIMINARY;      PRT;   292 AA.
AC   Q8BVU2;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   C-type lectin protein MT75 homolog.
GN   CHODL.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Head;
RX   MEDLINE=22354683; PubMed=12466851;
RA   The FANTOM Consortium,
RA   the RIKEN Genome Exploration Research Group Phase I & II Team;
RT   "Analysis of the mouse transcriptome based on functional annotation of
RT   60,770 full-length cDNAs.";
RL   Nature 420:563-573(2002).
DR   EMBL; AK076523; BAC36378.1; -.
DR   MGD; MGI:2179069; Chodl.
DR   GO; GO:0005529; F:sugar binding; IEA.
DR   InterPro; IPR001304; Lectin_C.
DR   Pfam; PF00059; lectin_c; 1.
DR   SMART; SM00034; CLECT; 1.
DR   PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
```

```
SQ   SEQUENCE   292 AA;  32502 MW;  73E631C0714D54E2 CRC64;

  Query Match             32.6%;  Score 652;  DB 11;  Length 292;
  Best Local Similarity   47.0%;  Pred. No. 9.4e-53;
  Matches  140;  Conservative   39;  Mismatches   79;  Indels   40;  Gaps    9;

Qy         10 VLLAVLLVGLRAATGRLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQ 69
              :| | ||    |   |::||| ||    : ||||: |||: | |::|:||: ||  :||
Db          8 LLGAALLCAQGAFARRVVSGQKVCFADVKHPCYKMAYFHELSSRVSFQEARLACESEGGV 67

Qy         70 LVSIESEDEQKLIEKFIENLLP-----SDGDFWIGLRRREEKQSNSTACQDLYAWTDGSI 124
              |:|:|:| |||||  ::||       ||||||||| |  : |: | || ||| |:|||
Db         68 LLSLENEAEQKLIESMLQNLTKPGTGISDGDFWIGLLRSGDGQT-SGACPDLYQWSDGSS 126

Qy        125 SQFRNWYVDEPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDE-KP 183
              ||||||| |||||||| ||||||||:|  |:||||::|||||||||:|:||||  |  |
Db        127 SQFRNWYTDEPSCGSEKCVVMYHQPTANPGLGGPYLYQWNDDRCNMKHNYICKYEPEIHP 186

Qy        184 AVPSREAEGEETELTTPVLPEETQEEDAKKTFKESREAAL--NLAYILIPSIPLLLLLVV 241
              |:    |:  ||   |||| |          || :  || |::||:|||||::|
Db        187 TEPA-----EKPYLTNQ--PEETHENVV------VTEAGIIPNLIYVIIPTIPLLLLILV 233

Qy        242 TTVVCWVWICRKRKR-----EQPDPSTKK-----------QHTIWPSPHQGNSPDL 281
                |   :  |||      :   | : :                ::  |   | ||| |
Db        234 ALGTCCFQMLHKRKARRHFIKDSTPLSSECLAESLNSNLVHMAGSLIPYHFQNNSPSL 291


RESULT 7
Q8BMI7
ID   Q8BMI7      PRELIMINARY;      PRT;   246 AA.
AC   Q8BMI7;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   C-type lectin protein MT75 homolog.
GN   CHODL.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Forelimb;
RX   MEDLINE=22354683; PubMed=12466851;
RA   The FANTOM Consortium,
RA   the RIKEN Genome Exploration Research Group Phase I & II Team;
RT   "Analysis of the mouse transcriptome based on functional annotation of
RT   60,770 full-length cDNAs.";
RL   Nature 420:563-573(2002).
DR   EMBL; AK031063; BAC27234.1; -.
DR   MGD; MGI:2179069; Chodl.
DR   GO; GO:0005529; F:sugar binding; IEA.
DR   InterPro; IPR001304; Lectin_C.
DR   Pfam; PF00059; lectin_c; 1.
DR   SMART; SM00034; CLECT; 1.
DR   PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
SQ   SEQUENCE   246 AA;  27394 MW;  E872660C58267752 CRC64;

  Query Match             32.3%;  Score 646;  DB 11;  Length 246;
  Best Local Similarity   52.2%;  Pred. No. 2.7e-52;
  Matches  132;  Conservative   34;  Mismatches   65;  Indels   22;  Gaps    7;

Qy         10 VLLAVLLVGLRAATGRLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQ 69
              :| | ||    |   |::||| ||    : ||||: |||: | |::|:||: ||  :||
Db          8 LLGAALLCAQGAFARRVVSGQKVCFADVKHPCYKMAYFHELSSRVSFQEARLACESEGGV 67

Qy         70 LVSIESEDEQKLIEKFIENLLP-----SDGDFWIGLRRREEKQSNSTACQDLYAWTDGSI 124
              |:|:|:| |||||  ::||       ||||||||| |  : |: | || ||| |:|||
Db         68 LLSLENEAEQKLIESMLQNLTKPGTGISDGDFWIGLLRSGDGQT-SGACPDLYQWSDGSS 126

Qy        125 SQFRNWYVDEPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDE-KP 183
              ||||||| |||||||| ||||||||:|  |:||||::|||||||||:|:||||  |  |
Db        127 SQFRNWYTDEPSCGSEKCVVMYHQPTANPGLGGPYLYQWNDDRCNMKHNYICKYEPEIHP 186

Qy        184 AVPSREAEGEETELTTPVLPEETQEEDAKKTFKESREAAL--NLAYILIPSIPLLLLLVV 241
              |:    |:  ||   |||| |          || :  || |::||:|||||::|
Db        187 TEPA-----EKPYLTNQ--PEETHENVV------VTEAGIIPNLIYVIIPTIPLLLLILV 233

Qy        242 TTVVCWVWICRKR 254
                |   :  ||
Db        234 ALGTCCFQMLHKR 246


RESULT 8
Q7Z799
ID   Q7Z799      PRELIMINARY;      PRT;   232 AA.
AC   Q7Z799;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Chondrolectin variant CHODLF.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22638322; PubMed=12621022;
RA   Weng L., Van Bockstaele D.R., Wauters J., Van Marck E., Plum J.,
RA   Berneman Z.N., Merregaert J.;
RT   "A Novel Alternative Spliced Chondrolectin Isoform Lacking the
RT   Transmembrane Domain Is Expressed during T Cell Maturation.";
RL   J. Biol. Chem. 278:19164-19170(2003).
DR   EMBL; AF523314; AAP43903.1; -.
SQ   SEQUENCE   232 AA;  25979 MW;  98B80BBDF7C7F439 CRC64;

  Query Match             28.8%;  Score 576;  DB 4;  Length 232;
  Best Local Similarity   49.0%;  Pred. No. 9.6e-46;
  Matches  121;  Conservative   35;  Mismatches   65;  Indels   26;  Gaps    9;

Qy         46 YFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLP-----SDGDFWIGL 100
              ||| | |::|:||: ||  :|| |:|:|:| ||||||  ::||       |||||||||
Db          3 YFHELSSRVSFQEARLACESEGGVLLSLENEAEQKLIESMLQNLTKPGTGISDGDFWIGL 62

Qy        101 RRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEPSCGSEVCVVMYHQPSAPAGIGGPYM 160
               |  : |: | || ||| |:||| ||:|||| |||||||| ||||||||:|  |:||||:
Db         63 WRNGDGQT-SGACPDLYQWSDGSNSQYRNWYTDEPSCGSEKCVVMYHQPTANPGLGGPYL 121

Qy        161 FQWNDDRCNMKNNFICKYSDE-KPAVPSREAEGEETELTTPVLPEETQEEDAKKTFKESR 219
              :|||||||||:|:||||  |  |  |      |:  ||   | :| :
Db        122 YQWNDDRCNMKHNYICKYEPEINPTAPV-----EKPYLTNQ--PGDTHQNVV------VT 168

Qy        220 EAAL--NLAYILIPSIPLLLLLVVTTVVCWVWICRKRK-REQPDPSTKKQHTIWPSPHQG 276
              || :  || |::||:||||||::|    |   :  | | | :  |:   | |:| |
Db        169 EAGIIPNLIYVVIPTIPLLLLILVAFGTCCFQMLHKSKGRTKTSPN---QSTLWISKSTR 225

Qy        277 NSPDLEV 283
                :||
Db        226 KESGMEV 232


RESULT 9
Q7Z798
ID   Q7Z798      PRELIMINARY;      PRT;   236 AA.
AC   Q7Z798;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Chondrolectin variant CHODLFdeltaE.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22638322; PubMed=12621022;
```

```
RA   Weng L., Van Bockstaele D.R., Wauters J., Van Marck E., Plum J.,
RA   Berneman Z.N., Merregaert J.;
RT   "A Novel Alternative Spliced Chondrolectin Isoform Lacking the
RT   Transmembrane Domain Is Expressed during T Cell Maturation.";
RL   J. Biol. Chem. 278:19164-19170(2003).
DR   EMBL; AF523315; AAP43904.1; -.
SQ   SEQUENCE   236 AA;  27185 MW;  5110CDC80C7C2CD9 CRC64;

  Query Match             25.7%;  Score 513;  DB 4;  Length 236;
  Best Local Similarity   54.3%;  Pred. No. 8.1e-40;
  Matches   94;  Conservative   28;  Mismatches   37;  Indels   14;  Gaps    3;

Qy         46 YFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLP-----SDGDFWIGL 100
              |||: | |::|:||: ||  :|| |:|:|:| ||||||  ::||       |||||||||
Db          3 YFHELSSRVSFQEARLACESEGGVLLSLENEAEQKLIESMLQNLTKPGTGISDGDFWIGL 62

Qy        101 RRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEPSCGSEVCVVMYHQPSAPAGIGGPYM 160
               |  : |: | || ||| |:||| ||:|||| |||||||| ||||||||:|  |:|||:
Db         63 WRNGDGQT-SGACPDLYQWSDGSNSQYRNWYTDEPSCGSEKCVVMYHQPTANPGLGGPYL 121

Qy        161 FQWNDDRCNMKNNFICKYSD--------EKPAVPSREAEGEETELTTPVLPEE 205
              :||||||||||:|:||||          ||| : ::  :  :  : |  : ||
Db        122 YQWNDDRCNMKHNYICKYEPEINPTAPVEKPYLTNQPGDTHQNVVVTEAVKEE 174


RESULT 10
Q7Z7A0
ID   Q7Z7A0      PRELIMINARY;      PRT;   236 AA.
AC   Q7Z7A0;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Chondrolectin variant CHODLdeltaE.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22638322; PubMed=12621022;
RA   Weng L., Van Bockstaele D.R., Wauters J., Van Marck E., Plum J.,
RA   Berneman Z.N., Merregaert J.;
RT   "A Novel Alternative Spliced Chondrolectin Isoform Lacking the
RT   Transmembrane Domain Is Expressed during T Cell Maturation.";
RL   J. Biol. Chem. 278:19164-19170(2003).
DR   EMBL; AF523313; AAP43902.1; -.
SQ   SEQUENCE   236 AA;  27115 MW;  9C6CE9A9EE7062E2 CRC64;

  Query Match             25.0%;  Score 500;  DB 4;  Length 236;
  Best Local Similarity   53.2%;  Pred. No. 1.4e-38;
  Matches   92;  Conservative   28;  Mismatches   39;  Indels   14;  Gaps    3;

Qy         46 YFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLP-----SDGDFWIGL 100
              |||: | |::|: |: ||  :|| |:|:|:| ||||||  ::||       |||||||||
Db          3 YFHELSSRVSFQGARLACESEGGVLLSLENEAEQKLIESMLQNLTKPGTGISDGDFWIGL 62

Qy        101 RRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEPSCGSEVCVVMYHQPSAPAGIGGPYM 160
               |  : |: | || ||| |:||| ||:|||| |||||||| ||||||||:|  |:| ||:
Db         63 WRNGDGQT-SGACPDLYQWSDGSNSQYRNWYTDEPSCGSEKCVVMYHQPTANPGLGSPYL 121

Qy        161 FQWNDDRCNMKNNFICKYSD--------EKPAVPSREAEGEETELTTPVLPEE 205
              :||||||||||:|:||||          ||| : ::  :  :  : |  : ||
Db        122 YQWNDDRCNMKHNYICKYEPEINPTAPVEKPYLTNQPGDTHQNVVVTEAVKEE 174


RESULT 11
Q9W6E1
ID   Q9W6E1      PRELIMINARY;      PRT;  1290 AA.
AC   Q9W6E1;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Neurocan core protein.
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
OX   NCBI_TaxID=9031;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20309833; PubMed=10851024;
RA   Li H., Leung T.C., Hoffman S., Balsamo J., Lilien J.;
RT   "Coordinate Regulation of Cadherin and Integrin Function by the
RT   Chondroitin Sulfate Proteoglycan Neurocan.";
RL   J. Cell Biol. 149:1275-1288(2000).
DR   EMBL; AF116856; AAD24546.2; -.
DR   HSSP; P08709; 1BF9.
DR   GO; GO:0005509; F:calcium ion binding; IEA.
DR   GO; GO:0005540; F:hyaluronic acid binding; IEA.
DR   GO; GO:0005529; F:sugar binding; IEA.
DR   GO; GO:0007155; P:cell adhesion; IEA.
DR   InterPro; IPR000152; Asx_hydroxyl_S.
DR   InterPro; IPR000742; EGF_2.
DR   InterPro; IPR001881; EGF_Ca.
DR   InterPro; IPR001438; EGF_II.
DR   InterPro; IPR006209; EGF_like.
DR   InterPro; IPR003599; Ig.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR001304; Lectin_C.
DR   InterPro; IPR000538; Link.
DR   InterPro; IPR000436; Sushi_SCR_CCP.
DR   Pfam; PF00008; EGF; 2.
DR   Pfam; PF00047; ig; 1.
DR   Pfam; PF00059; lectin_c; 1.
DR   Pfam; PF00084; sushi; 1.
DR   Pfam; PF00193; Xlink; 2.
DR   PRINTS; PR00010; EGFBLOOD.
DR   PRINTS; PR01265; LINKMODULE.
DR   ProDom; PD000918; Link; 2.
DR   SMART; SM00032; CCP; 1.
DR   SMART; SM00034; CLECT; 1.
DR   SMART; SM00179; EGF_CA; 1.
DR   SMART; SM00409; IG; 1.
DR   SMART; SM00445; LINK; 2.
DR   PROSITE; PS00010; ASX_HYDROXYL; 1.
DR   PROSITE; PS00615; C_TYPE_LECTIN_1; 1.
DR   PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
DR   PROSITE; PS00022; EGF_1; 2.
DR   PROSITE; PS01186; EGF_2; 1.
DR   PROSITE; PS01187; EGF_CA; 1.
DR   PROSITE; PS50835; IG_LIKE; 1.
DR   PROSITE; PS01241; LINK; 2.
KW   EGF-like domain.
SQ   SEQUENCE   1290 AA;  138877 MW;  182BD86D0E40BE78 CRC64;

  Query Match              9.2%;  Score 185;  DB 13;  Length 1290;
  Best Local Similarity   34.7%;  Pred. No. 5.2e-08;
  Matches   50;  Conservative   18;  Mismatches   42;  Indels   34;  Gaps    8;

Qy         38 QRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFW 97
              |  ||:  ||   ||| ::|:|:  |||  | | || |::|   |  |          : |
Db       1064 QGHCYR--YF---SRRRSWEDAERDCRRRAGHLTSIHSQEEHGFINSF------GHENTW 1112

Qy         98 IGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEPS---CGSEVCVVMY-HQPSAPA 153
              |||  |  :|          : ||| :  |: ||  ::|     | | |||:  |:
Db       1113 IGLNDRIVEQD--------FQWTDNTGLQYENWRENQPDNFFAGGEDCVVLVSHE----- 1159

Qy        154 GIGGPYMFQWNDDRCNMKNNFICK 177
               ||     :|||  ||    :|||
Db       1160 -IG-----KWNDVPCNYNLPYICK 1177


RESULT 12
Q61830
ID   Q61830      PRELIMINARY;      PRT;  1456 AA.
AC   Q61830;
```

```
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Macrophage mannose receptor precursor.
GN   MRC1.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6;
RX   MEDLINE=93043353; PubMed=1421407;
RA   Harris N., Rits M., Chang G., Ezekowitz R.B.;
RT   "Characterization of the murine macrophage mannose receptor.";
RL   Blood 80:2363-2373(1992).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6;
RA   Super M.;
RL   Submitted (APR-1992) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; Z11974; CAA78028.1; -.
DR   PIR; A48925; A48925.
DR   PDB; 1DQG; 10-MAY-00.
DR   PDB; 1DQO; 10-MAY-00.
DR   PDB; 1FWU; 17-JAN-01.
DR   PDB; 1FWV; 17-JAN-01.
DR   MGD; MGI:97142; Mrc1.
DR   GO; GO:0009928; C:cell surface (sensu Magnoliophyta); IDA.
DR   GO; GO:0005537; F:mannose binding; IDA.
DR   GO; GO:0004888; F:transmembrane receptor activity; IDA.
DR   InterPro; IPR002353; AntifreezeII.
DR   InterPro; IPR000562; FN_Type_II.
DR   InterPro; IPR001304; Lectin_C.
DR   InterPro; IPR008997; RicinB_like.
DR   InterPro; IPR000772; Ricin_B_lectin.
DR   Pfam; PF00040; fn2; 1.
DR   Pfam; PF00059; lectin_c; 8.
DR   Pfam; PF00652; Ricin_B_lectin; 2.
DR   PRINTS; PR00356; ANTIFREEZEII.
DR   PRINTS; PR00013; FNTYPEII.
DR   ProDom; PD000995; FN_Type_II; 1.
DR   SMART; SM00034; CLECT; 8.
DR   SMART; SM00059; FN2; 1.
DR   SMART; SM00458; RICIN; 1.
DR   PROSITE; PS00615; C_TYPE_LECTIN_1; 6.
DR   PROSITE; PS50041; C_TYPE_LECTIN_2; 8.
DR   PROSITE; PS00023; FIBRONECTIN_2; 1.
DR   PROSITE; PS50231; RICIN_B_LECTIN; 1.
KW   Receptor; Signal.
FT   SIGNAL       1     18       POTENTIAL.
FT   CHAIN       19   1456       MACROPHAGE MANNOSE RECEPTOR.
SQ   SEQUENCE   1456 AA;  165065 MW;  4EBD3F1B8619A594 CRC64;

  Query Match             8.9%;  Score 178.5;  DB 11;  Length 1456;
  Best Local Similarity   25.9%;  Pred. No. 2.5e-07;
  Matches   53;  Conservative   33;  Mismatches   66;  Indels   53;  Gaps    8;

Qy        42 YKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFWIGLR 101
             ||   :: :  :   : |:  |::: | | :|:|| |:| : |:| |       ::||:
Db       807 YKDYQYYFSKEKETMDNARRFCKKNFGDLATIKSESEKKFLWKYI-NKNGGQSPYFIGML 865

Qy       102 RREEKQSNSTACQDLYAWTDGSISQFRNWYVDEPSCGS--EVCVVMYHQPSAPAGIGGPY 159
             :|:         : | |||    |  |    ||:  :  | || ||      :|
Db       866 ISMDKK---------FIWMDGSKVDFVAWATGEPNFANDDENCVTMY----TNSGF---- 908

Qy       160 MFQWNDDRCNMKNNFICK---YSDEKPAVPSREAEGEETELTTPVLPEETQE-------- 208
             |||  |   |||||:    |    |:|          |||  |   :|
Db       909 ---WNDINCGYPNNFICQRHNSSINATAMP----------TTPTTPGGCKEGWHLYKNK 954

Qy       209 --------EDAKKTFKESREAALNL 225
                    : ||:::::|:|   |
Db       955 CFKIFGFANEEKKSWQDARQACKGL 979
```

```
RESULT 13
Q25199
ID   Q25199      PRELIMINARY;      PRT;  1348 AA.
AC   Q25199;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Tyrosine kinase receptor.
OS   Hydra attenuata (Hydra) (Hydra vulgaris).
OC   Eukaryota; Metazoa; Cnidaria; Hydrozoa; Hydroida; Anthomedusae;
OC   Hydridae; Hydra.
OX   NCBI_TaxID=6087;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Irvine;
RX   MEDLINE=20209407; PubMed=10744720;
RA   Reidling J.C., Miller M.A., Steele R.E.;
RT   "Sweet Tooth, a Novel Receptor Protein-tyrosine Kinase with C-type
RT   Lectin-like Extracellular Domains.";
RL   J. Biol. Chem. 275:10323-10330(2000).
DR   EMBL; L22612; AAA29218.2; -.
DR   HSSP; P11362; 1FGK.
DR   GO; GO:0005524; F:ATP binding; IEA.
DR   GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.
DR   GO; GO:0004872; F:receptor activity; IEA.
DR   GO; GO:0005529; F:sugar binding; IEA.
DR   GO; GO:0016740; F:transferase activity; IEA.
DR   GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR   InterPro; IPR001304; Lectin_C.
DR   InterPro; IPR003990; Pancreatis_ac.
DR   InterPro; IPR000719; Prot_kinase.
DR   InterPro; IPR001245; Tyr_pkinase.
DR   InterPro; IPR008266; Tyr_pkinase_AS.
DR   Pfam; PF00059; lectin_c; 4.
DR   Pfam; PF00069; pkinase; 1.
DR   PRINTS; PR01504; PNCREATITSAP.
DR   ProDom; PD000001; Prot_kinase; 2.
DR   SMART; SM00034; CLECT; 4.
DR   SMART; SM00219; TyrKc; 1.
DR   PROSITE; PS00615; C_TYPE_LECTIN_1; 2.
DR   PROSITE; PS50041; C_TYPE_LECTIN_2; 4.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR   PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
KW   ATP-binding; Kinase; Receptor; Transferase; Tyrosine-protein kinase.
SQ   SEQUENCE   1348 AA;  156916 MW;  77D2122093227FEF CRC64;

  Query Match             8.9%;  Score 177.5;  DB 5;  Length 1348;
  Best Local Similarity   28.0%;  Pred. No. 2.8e-07;
  Matches   47;  Conservative   32;  Mismatches   46;  Indels   43;  Gaps    8;

Qy        27 LSGQPVC--RGGTQRPC------YKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDE 78
             || : :|  :  |   |       |::  :   |   :: ::  :|:  || |:|||:::|
Db       412 LSHRFICKVKRATNEYCAEGWTSYRIYCYFIYSIEFDWFKSFSSCQNIGGNLLSIENQEE 471

Qy        79 QKLIEKFIENLLPSDGD-FWIGLRR------REEKQSNSTACQDLYAWTDGSISQFRNWY 131
                :|||| |  | | :|||| :      :: |:          : |:| : :|| ||
Db       472 ----NRFIENDLIKDNDKYWIGLNKIWNDYLKKNKR---------FEWSDNTYTQFFNWI 518

Qy       132 VDEP--SCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICK 177
              ::|  : | | || | :                |:|  | : | ||||
Db       519 TNQPDNNNGIESCVEMNYN------------GWSDKECKVLNGFICK 553


RESULT 14
Q9Y5P9
ID   Q9Y5P9      PRELIMINARY;      PRT;  1479 AA.
AC   Q9Y5P9;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
```

```
DE   Endocytic receptor Endo180.
GN   ENDO180.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20148849; PubMed=10683150;
RA   Sheikh H., Yarwood H., Ashworth A., Isacke C.;
RT   "Endo180, an endocytic recycling glycoprotein related to the
RT   macrophage mannose receptor is expressed on fibroblasts, endothelial
RT   cells and macrophages and functions as a lectin receptor.";
RL   J. Cell Sci. 113:1021-1032(2000).
DR   EMBL; AF134838; AAD30280.1; -.
DR   HSSP; P02751; 2FN2.
DR   GO; GO:0004872; F:receptor activity; IEA.
DR   GO; GO:0005529; F:sugar binding; IEA.
DR   GO; GO:0005215; F:transporter activity; IEA.
DR   GO; GO:0006810; P:transport; IEA.
DR   InterPro; IPR000562; FN_Type_II.
DR   InterPro; IPR001304; Lectin_C.
DR   InterPro; IPR000566; Lipocln_cytFABP.
DR   InterPro; IPR008997; RicinB_like.
DR   InterPro; IPR000772; Ricin_B_lectin.
DR   Pfam; PF00040; fn2; 1.
DR   Pfam; PF00059; lectin_c; 8.
DR   PRINTS; PR00013; FNTYPEII.
DR   ProDom; PD000995; FN_Type_II; 1.
DR   SMART; SM00034; CLECT; 8.
DR   SMART; SM00059; FN2; 1.
DR   SMART; SM00458; RICIN; 1.
DR   PROSITE; PS00615; C_TYPE_LECTIN_1; 3.
DR   PROSITE; PS50041; C_TYPE_LECTIN_2; 8.
DR   PROSITE; PS00023; FIBRONECTIN_2; 1.
DR   PROSITE; PS00213; LIPOCALIN; 1.
DR   PROSITE; PS50231; RICIN_B_LECTIN; 1.
KW   Receptor.
SQ   SEQUENCE   1479 AA;  166669 MW;  9F4BAF355F036FCE CRC64;

  Query Match             8.8%;  Score 176.5;  DB 4;  Length 1479;
  Best Local Similarity   30.4%;  Pred. No. 3.9e-07;
  Matches   51;  Conservative   25;  Mismatches   57;  Indels   35;  Gaps    7;

Qy         28 SGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIE 87
              | ||      |  ||::      : : :::|:|:|| | || |||| |  | : | | |:
Db        385 SWQPF-----QGHCYRL-----QAEKRSWQESKKACLRGGGDLVSIHSMAELEFITKQIK 434

Qy         88 NLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEPS---CGSEVCVV 144
                :    : ||||        |    |  : |:|||:  | :|:  ||:      | ||
Db        435 QEVE---ELWIGL--------NDLKLQMNFEWSDGSLVSFTHWHPFEPNNFRDSLEDCVT 483

Qy        145 MYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEG 192
              ::     | |       :|||  ||     ||| : :     : |  |
Db        484 IW----GPEG-------RWNDSPCNQSLPSICKKAGQLSQGAAEEDHG 520


RESULT 15
Q9UBG0
ID   Q9UBG0      PRELIMINARY;      PRT;  1479 AA.
AC   Q9UBG0;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Urokinase receptor-associated protein UPARAP.
GN   KIAA0709.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Behrendt N., Jensen O.N., Engelholm L.H., Mortz E., Mann M., Dano K.;
RT   "A urokinase receptor-associated protein with specific collagen-
RT   binding properties.";
RL   Submitted (NOV-1998) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Brain;
RX   MEDLINE=98403880; PubMed=9734811;
RA   Ishikawa K., Nagase T., Suyama M., Miyajima N., Tanaka A., Kotani H.,
RA   Nomura N., Ohara O.;
RT   "Prediction of the coding sequences of unidentified human genes. X.
RT   The complete sequences of 100 new cDNA clones from brain which can
RT   code for large proteins in vitro.";
RL   DNA Res. 5:169-176(1998).
DR   EMBL; AF107292; AAF14192.1; -.
DR   EMBL; AB014609; BAA31684.1; -.
DR   HSSP; P02751; 2FN2.
DR   Genew; HGNC:16875; MRC2.
DR   GO; GO:0016301; F:kinase activity; IEA.
DR   GO; GO:0004872; F:receptor activity; IEA.
DR   GO; GO:0005529; F:sugar binding; IEA.
DR   GO; GO:0005215; F:transporter activity; IEA.
DR   GO; GO:0006810; P:transport; IEA.
DR   InterPro; IPR000562; FN_Type_II.
DR   InterPro; IPR001304; Lectin_C.
DR   InterPro; IPR000566; Lipocln_cytFABP.
DR   InterPro; IPR008997; RicinB_like.
DR   InterPro; IPR000772; Ricin_B_lectin.
DR   Pfam; PF00040; fn2; 1.
DR   Pfam; PF00059; lectin_c; 8.
DR   PRINTS; PR00013; FNTYPEII.
DR   ProDom; PD000995; FN_Type_II; 1.
DR   SMART; SM00034; CLECT; 8.
DR   SMART; SM00059; FN2; 1.
DR   SMART; SM00458; RICIN; 1.
DR   PROSITE; PS00615; C_TYPE_LECTIN_1; 3.
DR   PROSITE; PS50041; C_TYPE_LECTIN_2; 8.
DR   PROSITE; PS00023; FIBRONECTIN_2; 1.
DR   PROSITE; PS00213; LIPOCALIN; 1.
DR   PROSITE; PS50231; RICIN_B_LECTIN; 1.
KW   Kinase; Receptor.
SQ   SEQUENCE   1479 AA;  166654 MW;  C7583EA78E2792D1 CRC64;

  Query Match             8.8%;  Score 176.5;  DB 4;  Length 1479;
  Best Local Similarity   30.4%;  Pred. No. 3.9e-07;
  Matches   51;  Conservative   25;  Mismatches   57;  Indels   35;  Gaps    7;

Qy         28 SGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIE 87
              | ||      |  ||::      : : :::|:|:|| | || |||| |  | : | | |:
Db        385 SWQPF-----QGHCYRL-----QAEKRSWQESKKACLRGGGDLVSIHSMAELEFITKQIK 434

Qy         88 NLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEPS---CGSEVCVV 144
                :    : ||||        |    |  : |:|||:  | :|:  ||:      | ||
Db        435 QEVE---ELWIGL--------NDLKLQMNFEWSDGSLVSFTHWHPFEPNNFRDSLEDCVT 483

Qy        145 MYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEG 192
              ::     | |       :|||  ||     ||| : :     : |  |
Db        484 IW----GPEG-------RWNDSPCNQSLPSICKKAGQLSQGAAEEDHG 520


Search completed: September  9, 2004, 22:22:34
Job time : 45 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: September 9, 2004, 22:20:26 ; Search time 34 Seconds
(without alignments)
567.886 Million cell updates/sec

Title: US-09-887-855-2
Perfect score: 2000
Sequence: 1 MRPGTALQAVLLAVLLVGLR.....PDQMGRSKESGWVENEIYGY 374

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1990	99.5	374	4	US-09-489-847-166	Sequence 166, App
2	1986	99.3	382	4	US-09-907-794A-137	Sequence 137, App
3	1986	99.3	382	4	US-09-905-125A-137	Sequence 137, App
4	1986	99.3	382	4	US-09-902-775A-137	Sequence 137, App
5	1198.5	59.9	260	4	US-09-638-203-3	Sequence 3, Appli
6	645	32.2	273	4	US-09-638-203-2	Sequence 2, Appli
7	336	16.8	81	4	US-09-489-847-325	Sequence 325, App
8	336	16.8	82	4	US-09-489-847-234	Sequence 234, App
9	336	16.8	115	4	US-09-489-847-324	Sequence 324, App
10	183	9.2	1456	4	US-09-976-594-168	Sequence 168, App
11	178.5	8.9	1455	3	US-08-840-062-5	Sequence 5, Appli
12	176.5	8.8	1479	3	US-08-840-062-4	Sequence 4, Appli
13	174	8.7	1257	1	US-08-340-428B-49	Sequence 49, Appl
14	174	8.7	2409	6	5180808-2	Patent No. 5180808
15	170.5	8.5	1479	3	US-08-840-062-2	Sequence 2, Appli
16	154.5	7.7	652	2	US-08-751-305-2	Sequence 2, Appli
17	152	7.6	912	5	PCT-US95-03747-2	Sequence 2, Appli
18	151	7.5	197	4	US-09-602-877A-99	Sequence 99, Appl
19	151	7.5	455	4	US-09-866-028-50	Sequence 50, Appl
20	149	7.4	174	1	US-07-641-971B-1	Sequence 1, Appli
21	149	7.4	174	1	US-07-781-248A-1	Sequence 1, Appli
22	149	7.4	320	1	US-08-365-103B-10	Sequence 10, Appl
23	149	7.4	321	1	US-08-365-103B-8	Sequence 8, Appli
24	148.5	7.4	404	4	US-09-517-605-2	Sequence 2, Appli
25	143.5	7.2	372	2	US-08-513-278-4	Sequence 4, Appli
26	143.5	7.2	372	6	5514582-4	Patent No. 5514582
27	142	7.1	1487	3	US-08-840-062-7	Sequence 7, Appli
28	141.5	7.1	125	3	US-08-722-126A-7	Sequence 7, Appli
29	141.5	7.1	125	5	PCT-US95-04258-7	Sequence 7, Appli
30	141.5	7.1	287	1	US-08-365-103B-4	Sequence 4, Appli
31	141.5	7.1	300	1	US-08-365-103B-6	Sequence 6, Appli
32	141.5	7.1	327	1	US-08-365-103B-2	Sequence 2, Appli
33	138.5	6.9	117	6	5514582-7	Patent No. 5514582
34	138.5	6.9	119	1	US-08-340-539A-12	Sequence 12, Appl
35	137.5	6.9	110	6	5514582-12	Patent No. 5514582
36	137.5	6.9	123	6	5514582-19	Patent No. 5514582
37	137.5	6.9	158	2	US-08-729-103-1	Sequence 1, Appli
38	137.5	6.9	158	2	US-08-468-413-2	Sequence 2, Appli
39	137.5	6.9	158	3	US-09-162-508-2	Sequence 2, Appli
40	137.5	6.9	158	5	PCT-US95-07169-2	Sequence 2, Appli
41	137.5	6.9	492	4	US-09-724-864-39	Sequence 39, Appl
42	135	6.8	238	3	US-09-111-470-8	Sequence 8, Appli
43	134.5	6.7	128	4	US-09-535-521-8	Sequence 8, Appli
44	134.5	6.7	139	4	US-09-535-521-11	Sequence 11, Appl
45	134.5	6.7	141	4	US-09-535-521-14	Sequence 14, Appl

ALIGNMENTS

RESULT 1
US-09-489-847-166
; Sequence 166, Application US/09489847
; Patent No. 6476195
; GENERAL INFORMATION:
; APPLICANT: Rosen et al
; TITLE OF INVENTION: 98 Human Secreted Proteins
; FILE REFERENCE: PZ031P1
; CURRENT APPLICATION NUMBER: US/09/489,847
; CURRENT FILING DATE: 2000-01-24
; EARLIER APPLICATION NUMBER: PCT/US99/17130
; EARLIER FILING DATE: 1999-07-29
; EARLIER APPLICATION NUMBER: 60/094,657
; EARLIER FILING DATE: 1998-07-30
; EARLIER APPLICATION NUMBER: 60/095,486
; EARLIER FILING DATE: 1998-08-05
; EARLIER APPLICATION NUMBER: 60/096,319
; EARLIER FILING DATE: 1998-08-12
; EARLIER APPLICATION NUMBER: 60/095,454
; EARLIER FILING DATE: 1998-08-06
; EARLIER APPLICATION NUMBER: 60/095,455
; EARLIER FILING DATE: 1998-08-06
; NUMBER OF SEQ ID NOS: 376
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 166
; LENGTH: 374
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (84)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (112)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-489-847-166

Query Match 99.5%; Score 1990; DB 4; Length 374;
Best Local Similarity 99.5%; Pred. No. 3.9e-193;
Matches 372; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy          1 MRPGTALQAVLLAVLLVGLRAATGRLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAK 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MRPGTALQAVLLAVLLVGLRAATGRLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAK 60

Qy         61 EACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWT 120
              |||||||||||||||||||||||| |||||||||||||||||||||||| ||||||||
Db         61 EACRRDGGQLVSIESEDEQKLIEXFIENLLPSDGDFWIGLRRREEKQSNSTXCQDLYAWT 120
```

```
Qy        121 DGSISQFRNWYVDEPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSD 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 DGSISQFRNWYVDEPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSD 180

Qy        181 EKPAVPSREAEGEETELTTPVLPEETQEEDAKKTFKESREAALNLAYILIPSIPLLLLLV 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 EKPAVPSREAEGEETELTTPVLPEETQEEDAKKTFKESREAALNLAYILIPSIPLLLLLV 240

Qy        241 VTTVVCWVWICRKRKREQPDPSTKKQHTIWPSPHQGNSPDLEVYNVIRKQSEADLAETRP 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 VTTVVCWVWICRKRKREQPDPSTKKQHTIWPSPHQGNSPDLEVYNVIRKQSEADLAETRP 300

Qy        301 DLKNISFRVCSGEATPDDMSCDYDNMAVNPSESGFVTLVSVESGFVTNDIYEFSPDQMGR 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 DLKNISFRVCSGEATPDDMSCDYDNMAVNPSESGFVTLVSVESGFVTNDIYEFSPDQMGR 360

Qy        361 SKESGWVENEIYGY 374
              ||||||||||||||
Db        361 SKESGWVENEIYGY 374
```

RESULT 2
US-09-907-794A-137
; Sequence 137, Application US/09907794A
; Patent No. 6635468
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/907,794A
; CURRENT FILING DATE: 2001-07-17
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 137
; LENGTH: 382
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-907-794A-137

Query Match 99.3%; Score 1986; DB 4; Length 382;
Best Local Similarity 97.9%; Pred. No. 1e-192;
Matches 374; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

```
Qy          1 MRPGTALQAVLLAVLLVGLRAATGRLLS--------GQPVCRGGTQRPCYKVIYFHDTSR 52
              |||||||||||||||||||||||||||        ||||||||||||||||||||||||
Db          1 MRPGTALQAVLLAVLLVGLRAATGRLLSASDLDLRGGQPVCRGGTQRPCYKVIYFHDTSR 60

Qy         53 RLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTA 112
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 RLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTA 120

Qy        113 CQDLYAWTDGSISQFRNWYVDEPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKN 172
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CQDLYAWTDGSISQFRNWYVDEPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKN 180

Qy        173 NFICKYSDEKPAVPSREAEGEETELTTPVLPEETQEEDAKKTFKESREAALNLAYILIPS 232
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 NFICKYSDEKPAVPSREAEGEETELTTPVLPEETQEEDAKKTFKESREAALNLAYILIPS 240

Qy        233 IPLLLLLVVTTVVCWVWICRKRKREQPDPSTKKQHTIWPSPHQGNSPDLEVYNVIRKQSE 292
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 IPLLLLLVVTTVVCWVWICRKRKREQPDPSTKKQHTIWPSPHQGNSPDLEVYNVIRKQSE 300

Qy        293 ADLAETRPDLKNISFRVCSGEATPDDMSCDYDNMAVNPSESGFVTLVSVESGFVTNDIYE 352
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ADLAETRPDLKNISFRVCSGEATPDDMSCDYDNMAVNPSESGFVTLVSVESGFVTNDIYE 360

Qy        353 FSPDQMGRSKESGWVENEIYGY 374
              ||||||||||||||||||||||
Db        361 FSPDQMGRSKESGWVENEIYGY 382
```

RESULT 3
US-09-905-125A-137
; Sequence 137, Application US/09905125A
; Patent No. 6664376
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.

; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/905,125A
; CURRENT FILING DATE: 2001-07-12
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 137
; LENGTH: 382
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-905-125A-137

Query Match 99.3%; Score 1986; DB 4; Length 382;
Best Local Similarity 97.9%; Pred. No. 1e-192;
Matches 374; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

```
Qy          1 MRPGTALQAVLLAVLLVGLRAATGRLLS--------GQPVCRGGTQRPCYKVIYFHDTSR 52
              ||||||||||||||||||||||||||||        ||||||||||||||||||||||||
Db          1 MRPGTALQAVLLAVLLVGLRAATGRLLSASDLDLRGGQPVCRGGTQRPCYKVIYFHDTSR 60

Qy         53 RLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTA 112
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 RLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTA 120

Qy        113 CQDLYAWTDGSISQFRNWYVDEPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKN 172
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CQDLYAWTDGSISQFRNWYVDEPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKN 180

Qy        173 NFICKYSDEKPAVPSREAEGEETELTTPVLPEETQEEDAKKTFKESREAALNLAYILIPS 232
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 NFICKYSDEKPAVPSREAEGEETELTTPVLPEETQEEDAKKTFKESREAALNLAYILIPS 240

Qy        233 IPLLLLLVVTTVVCWVWICRKRKREQPDPSTKKQHTIWPSPHQGNSPDLEVYNVIRKQSE 292
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 IPLLLLLVVTTVVCWVWICRKRKREQPDPSTKKQHTIWPSPHQGNSPDLEVYNVIRKQSE 300

Qy        293 ADLAETRPDLKNISFRVCSGEATPDDMSCDYDNMAVNPSESGFVTLVSVESGFVTNDIYE 352
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ADLAETRPDLKNISFRVCSGEATPDDMSCDYDNMAVNPSESGFVTLVSVESGFVTNDIYE 360

Qy        353 FSPDQMGRSKESGWVENEIYGY 374
              ||||||||||||||||||||||
Db        361 FSPDQMGRSKESGWVENEIYGY 382
```

RESULT 4
US-09-902-775A-137
; Sequence 137, Application US/09902775A
; Patent No. 6686451
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/902,775A
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05

```
;  PRIOR APPLICATION NUMBER: PCT/US99/28214
;  PRIOR FILING DATE: 1999-11-29
;  PRIOR APPLICATION NUMBER: PCT/US99/28313
;  PRIOR FILING DATE: 1999-11-30
;  PRIOR APPLICATION NUMBER: PCT/US99/28564
;  PRIOR FILING DATE: 1999-12-02
;  PRIOR APPLICATION NUMBER: PCT/US99/28565
;  PRIOR FILING DATE: 1999-12-02
;  PRIOR APPLICATION NUMBER: PCT/US99/30095
;  PRIOR FILING DATE: 1999-12-16
;  PRIOR APPLICATION NUMBER: PCT/US99/30911
;  PRIOR FILING DATE: 1999-12-20
;  PRIOR APPLICATION NUMBER: PCT/US99/30999
;  PRIOR FILING DATE: 1999-12-20
;  PRIOR APPLICATION NUMBER: PCT/US00/00219
;  PRIOR FILING DATE: 2000-01-05
;  NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 137
;   LENGTH: 382
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-902-775A-137

  Query Match             99.3%;  Score 1986;  DB 4;  Length 382;
  Best Local Similarity   97.9%;  Pred. No. 1e-192;
  Matches  374;  Conservative    0;  Mismatches    0;  Indels    8;  Gaps    1;

Qy          1 MRPGTALQAVLLAVLLVGLRAATGRLLS--------GQPVCRGGTQRPCYKVIYFHDTSR 52
              |||||||||||||||||||||||||||        ||||||||||||||||||||||||
Db          1 MRPGTALQAVLLAVLLVGLRAATGRLLSASDLDLRGGQPVCRGGTQRPCYKVIYFHDTSR 60

Qy         53 RLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTA 112
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 RLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTA 120

Qy        113 CQDLYAWTDGSISQFRNWYVDEPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKN 172
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CQDLYAWTDGSISQFRNWYVDEPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKN 180

Qy        173 NFICKYSDEKPAVPSREAEGEETELTTPVLPEETQEEDAKKTFKESREAALNLAYILIPS 232
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 NFICKYSDEKPAVPSREAEGEETELTTPVLPEETQEEDAKKTFKESREAALNLAYILIPS 240

Qy        233 IPLLLLLVVTTVVCWVWICRKRKREQPDPSTKKQHTIWPSPHQGNSPDLEVYNVIRKQSE 292
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 IPLLLLLVVTTVVCWVWICRKRKREQPDPSTKKQHTIWPSPHQGNSPDLEVYNVIRKQSE 300

Qy        293 ADLAETRPDLKNISFRVCSGEATPDDMSCDYDNMAVNPSESGFVTLVSVESGFVTNDIYE 352
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ADLAETRPDLKNISFRVCSGEATPDDMSCDYDNMAVNPSESGFVTLVSVESGFVTNDIYE 360

Qy        353 FSPDQMGRSKESGWVENEIYGY 374
              ||||||||||||||||||||||
Db        361 FSPDQMGRSKESGWVENEIYGY 382


RESULT 5
US-09-638-203-3
; Sequence 3, Application US/09638203
; Patent No. 6602501
; GENERAL INFORMATION:
;  APPLICANT:  Daniel E.H. Afar
;  APPLICANT:  Rene S. Hubert
;  APPLICANT:  Aya Jakobovits
;  APPLICANT:  Arthur B. Raitano
;  TITLE OF INVENTION:  NOVEL C-TYPE LECTIN TRANSMEMBRANE
;  TITLE OF INVENTION:  ANTIGEN EXPRESSED IN HUMAN PROSTATE CANCER AND USES THEREOF
;  FILE REFERENCE:  129.20USU1
;  CURRENT APPLICATION NUMBER: US/09/638,203
;  CURRENT FILING DATE:  2000-08-11
;  PRIOR APPLICATION NUMBER:  60/148,935
;  PRIOR FILING DATE:  1999-08-12
;  NUMBER OF SEQ ID NOS:  47
;  SOFTWARE:  FastSEQ for Windows Version 4.0
; SEQ ID NO 3
;   LENGTH: 260
;   TYPE: PRT
;   ORGANISM: Hamster
US-09-638-203-3

  Query Match             59.9%;  Score 1198.5;  DB 4;  Length 260;
  Best Local Similarity   84.2%;  Pred. No. 3.8e-113;
  Matches  219;  Conservative   17;  Mismatches   23;  Indels    1;  Gaps    1;

Qy         25 RLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEK 84
              ||||||  ||||||:|||||||||||  :|||||||||||||||||||||:||||:||||
Db          1 RLLSGQLVCRGGTRRPCYKVIYFHDAFQRLNFEEAKEACRRDGGQLVSIETEDEQRLIEK 60

Qy         85 FIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEPSCGSEVCVV 144
              ||||||  ||||||||||| | || |:|||||||||||||  ||||||||||||||||||
Db         61 FIENLLASDGDFWIGLRRLEVKQVNNTACQDLYAWTDGSTSQFRNWYVDEPSCGSEVCVV 120

Qy        145 MYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPA-VPSREAEGEETELTTPVLP 203
              |||||||| |||| |||||||||||||||||||:||||:  ||    || ||  |||||
Db        121 MYHQPSAPPGIGGSYMFQWNDDRCNMKNNFICKYADEKPSTTPSIRPGGEATEPPTPVLP 180

Qy        204 EETQEEDAKKTFKESREAALNLAYILIPSIPLLLLLVVTTVVCWVWICRKRKREQPDPST 263
              ||||:|| |:|||||||||||||||||||||||| ||||||:  |||||||:||:|||||:|
Db        181 EETQKEDTKETFKESREAALNLAYILIPSIPLFLLLVVTSAACWVWICRRRKQEQPDPTT 240

Qy        264 KKQHTIWPSPHQGNSPDLEV 283
              |:|||||:||| |||:|:|
Db        241 KEQHTIWPTPHQENSPNLDV 260


RESULT 6
US-09-638-203-2
; Sequence 2, Application US/09638203
; Patent No. 6602501
; GENERAL INFORMATION:
;  APPLICANT:  Daniel E.H. Afar
;  APPLICANT:  Rene S. Hubert
;  APPLICANT:  Aya Jakobovits
;  APPLICANT:  Arthur B. Raitano
;  TITLE OF INVENTION:  NOVEL C-TYPE LECTIN TRANSMEMBRANE
;  TITLE OF INVENTION:  ANTIGEN EXPRESSED IN HUMAN PROSTATE CANCER AND USES THEREOF
;  FILE REFERENCE:  129.20USU1
;  CURRENT APPLICATION NUMBER: US/09/638,203
;  CURRENT FILING DATE:  2000-08-11
;  PRIOR APPLICATION NUMBER:  60/148,935
;  PRIOR FILING DATE:  1999-08-12
;  NUMBER OF SEQ ID NOS:  47
;  SOFTWARE:  FastSEQ for Windows Version 4.0
; SEQ ID NO 2
;   LENGTH: 273
;   TYPE: PRT
;   ORGANISM: Homo Sapiens
US-09-638-203-2

  Query Match             32.2%;  Score 645;  DB 4;  Length 273;
  Best Local Similarity   48.4%;  Pred. No. 5.1e-57;
  Matches  137;  Conservative   40;  Mismatches   80;  Indels   26;  Gaps    9;

Qy         10 VLLAVLLVGLRAATGRLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQ 69
              :| | || |  |   |::||| ||    : ||||: |||: | |::|:||: ||  :||
Db          8 LLGAALLCGHGAFCRRVVSGQKVCFADPKHPCYKMAYFHELSSRVSFQEARLACESEGGV 67

Qy         70 LVSIESEDEQKLIEKFIENLLP-----SDGDFWIGLRRREEKQSNSTACQDLYAWTDGSI 124
              |:|:|:| ||||||  ::||       |||||||| |  : |: | || ||| |:|||
Db         68 LLSLENEAEQKLIESMLQNLTKPGTGISDGDFWIGLWRNGDGQT-SGACPDLYQWSDGSN 126

Qy        125 SQFRNWYVDEPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDE-KP 183
              ||:|||| ||||||||.||||||||:|  |:||||::|||||||||:|:||||  |  |
Db        127 SQYRNWYTDEPSCGSEKCVVMYHQPTANPGLGGPYLYQWNDDRCNMKHNYICKYEPEINP 186

Qy        184 AVPSREAEGEETELTTPVLPEETQEEDAKKTFKESREAAL--NLAYILIPSIPLLLLLVV 241
```

```
                |      |:  ||    | :| :           || :  || |::||:||||||::|
Db        187 TAPV-----EKPYLTNQ--PGDTHQNVV------VTEAGIIPNLIYVVIPTIPLLLLILV 233

Qy        242 TTVVCWVVWICRKRK-REQPDPSTKKQHTIWPSPHQGNSPDLEV 283
                |   :  | | | :  |:   | |:| |         :||
Db        234 AFGTCCFQMLHKSKGRTKTSPN---QSTLWISKSTRKESGMEV 273


RESULT 7
US-09-489-847-325
; Sequence 325, Application US/09489847
; Patent No. 6476195
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al
;  TITLE OF INVENTION: 98 Human Secreted Proteins
;  FILE REFERENCE: PZ031P1
;  CURRENT APPLICATION NUMBER: US/09/489,847
;  CURRENT FILING DATE: 2000-01-24
;  EARLIER APPLICATION NUMBER: PCT/US99/17130
;  EARLIER FILING DATE: 1999-07-29
;  EARLIER APPLICATION NUMBER: 60/094,657
;  EARLIER FILING DATE: 1998-07-30
;  EARLIER APPLICATION NUMBER: 60/095,486
;  EARLIER FILING DATE: 1998-08-05
;  EARLIER APPLICATION NUMBER: 60/096,319
;  EARLIER FILING DATE: 1998-08-12
;  EARLIER APPLICATION NUMBER: 60/095,454
;  EARLIER FILING DATE: 1998-08-06
;  EARLIER APPLICATION NUMBER: 60/095,455
;  EARLIER FILING DATE: 1998-08-06
;  NUMBER OF SEQ ID NOS: 376
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 325
;   LENGTH: 81
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-489-847-325

  Query Match             16.8%;  Score 336;  DB 4;  Length 81;
  Best Local Similarity   100.0%;  Pred. No. 1.7e-26;
  Matches   65;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MRPGTALQAVLLAVLLVGLRAATGRLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAK 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MRPGTALQAVLLAVLLVGLRAATGRLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAK 60

Qy         61 EACRR 65
              |||||
Db         61 EACRR 65


RESULT 8
US-09-489-847-234
; Sequence 234, Application US/09489847
; Patent No. 6476195
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al
;  TITLE OF INVENTION: 98 Human Secreted Proteins
;  FILE REFERENCE: PZ031P1
;  CURRENT APPLICATION NUMBER: US/09/489,847
;  CURRENT FILING DATE: 2000-01-24
;  EARLIER APPLICATION NUMBER: PCT/US99/17130
;  EARLIER FILING DATE: 1999-07-29
;  EARLIER APPLICATION NUMBER: 60/094,657
;  EARLIER FILING DATE: 1998-07-30
;  EARLIER APPLICATION NUMBER: 60/095,486
;  EARLIER FILING DATE: 1998-08-05
;  EARLIER APPLICATION NUMBER: 60/096,319
;  EARLIER FILING DATE: 1998-08-12
;  EARLIER APPLICATION NUMBER: 60/095,454
;  EARLIER FILING DATE: 1998-08-06
;  EARLIER APPLICATION NUMBER: 60/095,455
;  EARLIER FILING DATE: 1998-08-06
;  NUMBER OF SEQ ID NOS: 376
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 234
;   LENGTH: 82
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: SITE
;   LOCATION: (82)
;   OTHER INFORMATION: Xaa equals stop translation
US-09-489-847-234

  Query Match             16.8%;  Score 336;  DB 4;  Length 82;
  Best Local Similarity   100.0%;  Pred. No. 1.7e-26;
  Matches   65;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MRPGTALQAVLLAVLLVGLRAATGRLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAK 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MRPGTALQAVLLAVLLVGLRAATGRLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAK 60

Qy         61 EACRR 65
              |||||
Db         61 EACRR 65


RESULT 9
US-09-489-847-324
; Sequence 324, Application US/09489847
; Patent No. 6476195
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al
;  TITLE OF INVENTION: 98 Human Secreted Proteins
;  FILE REFERENCE: PZ031P1
;  CURRENT APPLICATION NUMBER: US/09/489,847
;  CURRENT FILING DATE: 2000-01-24
;  EARLIER APPLICATION NUMBER: PCT/US99/17130
;  EARLIER FILING DATE: 1999-07-29
;  EARLIER APPLICATION NUMBER: 60/094,657
;  EARLIER FILING DATE: 1998-07-30
;  EARLIER APPLICATION NUMBER: 60/095,486
;  EARLIER FILING DATE: 1998-08-05
;  EARLIER APPLICATION NUMBER: 60/096,319
;  EARLIER FILING DATE: 1998-08-12
;  EARLIER APPLICATION NUMBER: 60/095,454
;  EARLIER FILING DATE: 1998-08-06
;  EARLIER APPLICATION NUMBER: 60/095,455
;  EARLIER FILING DATE: 1998-08-06
;  NUMBER OF SEQ ID NOS: 376
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 324
;   LENGTH: 115
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-489-847-324

  Query Match             16.8%;  Score 336;  DB 4;  Length 115;
  Best Local Similarity   100.0%;  Pred. No. 2.8e-26;
  Matches   65;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MRPGTALQAVLLAVLLVGLRAATGRLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAK 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         35 MRPGTALQAVLLAVLLVGLRAATGRLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAK 94

Qy         61 EACRR 65
              |||||
Db         95 EACRR 99


RESULT 10
US-09-976-594-168
; Sequence 168, Application US/09976594
; Patent No. 6673549
; GENERAL INFORMATION:
```

```
;  APPLICANT: Furness, Michael
;  APPLICANT:  Buchbinder, Jenny
;  TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROIDS
;  FILE REFERENCE: PA-0041 US
;  CURRENT APPLICATION NUMBER: US/09/976,594
;  CURRENT FILING DATE:  2001-10-12
;  PRIOR APPLICATION NUMBER: 60/240,409
;  PRIOR FILING DATE: 2000-10-12
;  NUMBER OF SEQ ID NOS: 1143
;  SOFTWARE: PERL Program
; SEQ ID NO 168
;   LENGTH: 1456
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No. 6673549 1674368CD1
US-09-976-594-168

  Query Match             9.2%;  Score 183;  DB 4;  Length 1456;
  Best Local Similarity   25.1%;  Pred. No. 4.4e-09;
  Matches   53;  Conservative   38;  Mismatches   66;  Indels   54;  Gaps    8;

Qy         42 YKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFWIGLR 101
              ||   :: :  :   : |:  |:|: | ||||:|| |:| : |:: |   :   ::|||
Db        807 YKDYQYYFSKEKETMDNARAFCKRNFGDLVSIQSESEKKFLWKYV-NRNDAQSAYFIGLL 865

Qy        102 RREEKQSNSTACQDLYAWTDGSISQFRNWYVDEPSCGS--EVCVVMYHQPSAPAGIGGPY 159
                 :|:          :|| |||    : :|   ||:  :  | || ||     : :|
Db        866 ISLDKK---------FAWMDGSKVDYVSWATGEPNFANEDENCVTMY----SNSGF---- 908

Qy        160 MFQWNDDRCNMKNNFICKYSDEK----PAVPSREAEGEETELTTPVLPEETQE------- 208
                 |||  |   | |||:  :         :|           || :|   :|
Db        909 ---WNDINCGYPNAFICQRHNSSINATTVMP----------TMPSVPSGCKEGWNFYSN 954

Qy        209 ---------EDAKKTFKESREAALNLAYILI 230
                      |: :| ::|:|:| :     |:
Db        955 KCFKIFGFMEEERKNWQEARKACIGFGGNLV 985


RESULT 11
US-08-840-062-5
; Sequence 5, Application US/08840062
; Patent No. 6117977
;  GENERAL INFORMATION:
;    APPLICANT:  LASKY, LAURENCE A.
;    APPLICANT:  WU, KAI
;    TITLE OF INVENTION:  TYPE C LECTINS
;    NUMBER OF SEQUENCES:  15
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Genentech, Inc.
;      STREET:  460 Point San Bruno Blvd
;      CITY:  South San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94080
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5 inch, 1.44 Mb floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  WinPatin (Genentech)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/840,062
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Dreger, Ginger R.
;      REGISTRATION NUMBER:  33,055
;      REFERENCE/DOCKET NUMBER:  P1019R1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415/225-3216
;      TELEFAX:  415/952-9881
;      TELEX:  910/371-7168
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1455 amino acids
;      TYPE:  Amino Acid
;      TOPOLOGY:  Linear
US-08-840-062-5

  Query Match             8.9%;  Score 178.5;  DB 3;  Length 1455;
  Best Local Similarity   25.9%;  Pred. No. 1.3e-08;
  Matches   53;  Conservative   33;  Mismatches   66;  Indels   53;  Gaps    8;

Qy         42 YKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFWIGLR 101
              ||   :: :  :   : |:  |:|: | ||||:|| |:| : |:: |   :   ::|||
Db        806 YKDYQYYFSKEKETMDNARRFCKKNFGDLATIKSESEKKFLWKYI-NKNGGQSPYFIGML 864

Qy        102 RREEKQSNSTACQDLYAWTDGSISQFRNWYVDEPSCGS--EVCVVMYHQPSAPAGIGGPY 159
                 :|:          : | |||    |  |    ||:  :  | || ||     :|
Db        865 ISMDKK---------FIWMDGSKVDFVAWATGEPNFANDDENCVTMY----TNSGF---- 907

Qy        160 MFQWNDDRCNMKNNFICK---YSDEKPAVPSREAEGEETELTTPVLPEETQE-------- 208
                 |||  |   |||||:    |    |:|           |||  |   :|
Db        908 ---WNDINCGYPNNFICQRHNSSINATAMP----------TTPTTPGGCKEGWHLYKNK 953

Qy        209 --------EDAKKTFKESREAALNL 225
                      : ||:::::|:|    |
Db        954 CFKIFGFANEEKKSWQDARQACKGL 978


RESULT 12
US-08-840-062-4
; Sequence 4, Application US/08840062
; Patent No. 6117977
;  GENERAL INFORMATION:
;    APPLICANT:  LASKY, LAURENCE A.
;    APPLICANT:  WU, KAI
;    TITLE OF INVENTION:  TYPE C LECTINS
;    NUMBER OF SEQUENCES:  15
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Genentech, Inc.
;      STREET:  460 Point San Bruno Blvd
;      CITY:  South San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94080
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5 inch, 1.44 Mb floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  WinPatin (Genentech)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/840,062
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Dreger, Ginger R.
;      REGISTRATION NUMBER:  33,055
;      REFERENCE/DOCKET NUMBER:  P1019R1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415/225-3216
;      TELEFAX:  415/952-9881
;      TELEX:  910/371-7168
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1479 amino acids
;      TYPE:  Amino Acid
;      TOPOLOGY:  Linear
US-08-840-062-4

  Query Match             8.8%;  Score 176.5;  DB 3;  Length 1479;
  Best Local Similarity   30.4%;  Pred. No. 2.1e-08;
  Matches   51;  Conservative   25;  Mismatches   57;  Indels   35;  Gaps    7;

Qy         28 SGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIE 87
```

```
Db        385 SWQPF-----QGHCYRL-----QAEKRSWQESKKACLRGGGDLVSIHSMAELEFITKQIK 434

Qy         88 NLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEPS---CGSEVCVV 144

Db        435 QEVE---ELWIGL--------NDLKLQMNFEWSDGSLVSFTHWHPFEPNNFRDSLEDCVT 483

Qy        145 MYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEG 192

Db        484 IW----GPEG-------RWNDSPCNQSLPSICKKAGQLSQGAAEEDHG 520


RESULT 13
US-08-340-428B-49
; Sequence 49, Application US/08340428B
; Patent No. 5648465
;  GENERAL INFORMATION:
;    APPLICANT:  MARGOLIS, Richard U.
;    APPLICANT:  RAUCH, Uwe
;    APPLICANT:  MARGOLIS, Renee K.
;    TITLE OF INVENTION:  CLONING, EXPRESSION AND USES FOR A
;    TITLE OF INVENTION:  NEUROCAN AS A CHONDROITIN SULFATE PROTEOGLYCAN
;    NUMBER OF SEQUENCES:  49
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Browdy and Neimark
;      STREET:  419 Seventh Street, N.W.
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  U.S.A.
;      ZIP:  20004
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/340,428B
;      FILING DATE:  14 No. 5648465ember 1994
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/922,911
;      FILING DATE:  03 August 1992
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Browdy, Roger L.
;      REGISTRATION NUMBER:  25,618
;      REFERENCE/DOCKET NUMBER:  Margolis=1A
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202-628-5197
;      TELEFAX:  202-737-3528
;  INFORMATION FOR SEQ ID NO:  49:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1257 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-340-428B-49

  Query Match             8.7%;  Score 174;  DB 1;  Length 1257;
  Best Local Similarity   30.8%;  Pred. No. 2.9e-08;
  Matches   44;  Conservative   18;  Mismatches   49;  Indels   32;  Gaps    6;

Qy          38 QRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFW 97

Db        1037 QGHCYR--YF---AHRRAWEDAERDCRRRAGHLTSVHSPEEHKFINSF------GHENSW 1085

Qy          98 IGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEPS---CGSEVCVVMYHQPSAPAG 154

Db        1086 IGLNDRTVERD--------FQWTDNTGLQYENWREKQPDNFFAGGEDCVVMVAHENG--- 1134

Qy         155 IGGPYMFQWNDDRCNMKNNFICK 177

Db        1135 -------RWNDVPCNYNLPYVCK 1150


RESULT 14
5180808-2
;Patent No. 5180808
;    APPLICANT: RUOSLAHTI, ERKKI I.
;    TITLE OF INVENTION: VERSICAN CORE PROTEIN, NUCLEIC ACID
;SEQUENCES ENCODING THE SAME, NUCLEIC ACID PROBES, ANTI-VERSICAN
;ANTIBODIES, AND METHODS OF DETECTING THE SAME
;    NUMBER OF SEQUENCES: 4
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/441,179
;      FILING DATE: 27-NOV-1989
;SEQ ID NO:2:
;      LENGTH: 2409
5180808-2

  Query Match             8.7%;  Score 174;  DB 6;  Length 2409;
  Best Local Similarity   28.5%;  Pred. No. 7.8e-08;
  Matches   47;  Conservative   23;  Mismatches   55;  Indels   40;  Gaps    8;

Qy          38 QRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDF- 96

Db        2190 QGQCYK--YF---AHRRTWDAAERECRLQGAHLTSILSHEEQMFVNRV-------GHDYQ 2237

Qy          97 WIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEP----SCGSEVCVVMYHQPSAP 152

Db        2238 WIGL--------NDKMFEHDFRWTDGSTLQYENWRPNQPDSFFSAGEDCVVIIWHENG-- 2287

Qy         153 AGIGGPYMFQWNDDRCNMKNNFICKYS----DEKPAVPSREAEGE 193

Db        2288 ---------QWNDVPCNYHLTYTCKKGTVACGQPPVVENAKTFGK 2323


RESULT 15
US-08-840-062-2
; Sequence 2, Application US/08840062
; Patent No. 6117977
;  GENERAL INFORMATION:
;    APPLICANT:  LASKY, LAURENCE A.
;    APPLICANT:  WU, KAI
;    TITLE OF INVENTION:  TYPE C LECTINS
;    NUMBER OF SEQUENCES:  15
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Genentech, Inc.
;      STREET:  460 Point San Bruno Blvd
;      CITY:  South San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94080
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5 inch, 1.44 Mb floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  WinPatin (Genentech)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/840,062
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Dreger, Ginger R.
;      REGISTRATION NUMBER:  33,055
;      REFERENCE/DOCKET NUMBER:  P1019R1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415/225-3216
;      TELEFAX:  415/952-9881
;      TELEX:  910/371-7168
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1479 amino acids
;      TYPE:  Amino Acid
;      TOPOLOGY:  Linear
```

US-08-840-062-2

Query Match 8.5%; Score 170.5; DB 3; Length 1479;
Best Local Similarity 31.4%; Pred. No. 8.4e-08;
Matches 48; Conservative 22; Mismatches 48; Indels 35; Gaps 7;

```
Qy          28 SGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIE 87
               | ||       |  ||::      : : :::|:| || | || |:|| |  | : | | |:
Db         384 SWQPF-----QGHCYRL-----QAEKRSWQESKRACLRGGGDLLSIHSMAELEFITKQIK 433

Qy          88 NLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEPS---CGSEVCVV 144
                 :    : ||||        |    |  : |:|||:  | :|:  ||:       | ||
Db         434 QEVE---ELWIGL--------NDLKLQMNFEWSDGSLVSFTHWHPFEPNNFRDSLEDCVT 482

Qy         145 MYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICK 177
               ::     | |       :|||  ||     |||
Db         483 IW----GPEG-------RWNDSPCNQSLPSICK 504
```

Search completed: September 9, 2004, 22:23:40
Job time : 35 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: September 9, 2004, 22:22:42 ; Search time 130 Seconds
(without alignments)
922.599 Million cell updates/sec

Title: US-09-887-855-2
Perfect score: 2000
Sequence: 1 MRPGTALQAVLLAVLLVGLR.....PDQMGRSKESGWVENEIYGY 374

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1335176 seqs, 320689617 residues

Total number of hits satisfying chosen parameters: 1335176

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2000	100.0	374	9	US-09-887-855-2	Sequence 2, Appli
2	2000	100.0	374	14	US-10-149-819-15	Sequence 15, Appl
3	2000	100.0	374	15	US-10-094-749-2090	Sequence 2090, Ap
4	2000	100.0	374	15	US-10-264-237-2579	Sequence 2579, Ap
5	2000	100.0	374	16	US-10-648-593-148	Sequence 148, App
6	1996	99.8	374	15	US-10-094-749-2142	Sequence 2142, Ap
7	1990	99.5	374	12	US-10-351-334-166	Sequence 166, App
8	1986	99.3	382	9	US-09-909-320-137	Sequence 137, App
9	1986	99.3	382	9	US-09-909-088B-137	Sequence 137, App
10	1986	99.3	382	9	US-09-905-291A-137	Sequence 137, App
11	1986	99.3	382	9	US-09-902-853-137	Sequence 137, App
12	1986	99.3	382	9	US-09-907-824-137	Sequence 137, App
13	1986	99.3	382	9	US-09-907-841-137	Sequence 137, App
14	1986	99.3	382	10	US-09-904-011-137	Sequence 137, App
15	1986	99.3	382	10	US-09-906-742-137	Sequence 137, App
16	1986	99.3	382	10	US-09-906-838-137	Sequence 137, App
17	1986	99.3	382	10	US-09-907-613-137	Sequence 137, App
18	1986	99.3	382	10	US-09-907-942-137	Sequence 137, App
19	1986	99.3	382	10	US-09-904-859-137	Sequence 137, App
20	1986	99.3	382	10	US-09-909-204-137	Sequence 137, App
21	1986	99.3	382	10	US-09-904-820-137	Sequence 137, App
22	1986	99.3	382	10	US-09-904-786-137	Sequence 137, App
23	1986	99.3	382	10	US-09-906-646-137	Sequence 137, App
24	1986	99.3	382	10	US-09-906-700-137	Sequence 137, App
25	1986	99.3	382	10	US-09-903-786-137	Sequence 137, App
26	1986	99.3	382	10	US-09-902-903-137	Sequence 137, App
27	1986	99.3	382	10	US-09-903-749A-137	Sequence 137, App
28	1986	99.3	382	10	US-09-904-119-137	Sequence 137, App
29	1986	99.3	382	10	US-09-904-956-137	Sequence 137, App
30	1986	99.3	382	10	US-09-902-736-137	Sequence 137, App
31	1986	99.3	382	10	US-09-907-794-137	Sequence 137, App
32	1986	99.3	382	10	US-09-903-943-137	Sequence 137, App
33	1986	99.3	382	10	US-09-904-462-137	Sequence 137, App
34	1986	99.3	382	10	US-09-907-925-137	Sequence 137, App
35	1986	99.3	382	10	US-09-902-692-137	Sequence 137, App
36	1986	99.3	382	10	US-09-903-520-137	Sequence 137, App
37	1986	99.3	382	10	US-09-905-056-137	Sequence 137, App
38	1986	99.3	382	10	US-09-909-064-137	Sequence 137, App
39	1986	99.3	382	10	US-09-904-553-137	Sequence 137, App
40	1986	99.3	382	10	US-09-905-381-137	Sequence 137, App
41	1986	99.3	382	10	US-09-905-088-137	Sequence 137, App
42	1986	99.3	382	10	US-09-907-575-137	Sequence 137, App
43	1986	99.3	382	10	US-09-905-075-137	Sequence 137, App
44	1986	99.3	382	10	US-09-902-759-137	Sequence 137, App
45	1986	99.3	382	10	US-09-902-634-137	Sequence 137, App

ALIGNMENTS

RESULT 1
US-09-887-855-2
; Sequence 2, Application US/09887855
; Patent No. US20020058310A1
; GENERAL INFORMATION:
; APPLICANT: Immunex Corporation
; APPLICANT: Anderson, Dirk M
; TITLE OF INVENTION: LECTIN SS3939 DNA AND POLYPEPTIDES
; FILE REFERENCE: 2883-US
; CURRENT APPLICATION NUMBER: US/09/887,855
; CURRENT FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 374
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-887-855-2

Query Match 100.0%; Score 2000; DB 9; Length 374;
Best Local Similarity 100.0%; Pred. No. 2.8e-179;
Matches 374; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MRPGTALQAVLLAVLLVGLRAATGRLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAK 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MRPGTALQAVLLAVLLVGLRAATGRLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAK 60

Qy         61 EACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 EACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWT 120

Qy        121 DGSISQFRNWYVDEPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSD 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 DGSISQFRNWYVDEPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSD 180

Qy        181 EKPAVPSREAEGEETELTTPVLPEETQEEDAKKTFKESREAALNLAYILIPSIPLLLLLV 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 EKPAVPSREAEGEETELTTPVLPEETQEEDAKKTFKESREAALNLAYILIPSIPLLLLLV 240
```

```
Qy        241 VTTVVCWVWICRKRKREQPDPSTKKQHTIWPSPHQGNSPDLEVYNVIRKQSEADLAETRP 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 VTTVVCWVWICRKRKREQPDPSTKKQHTIWPSPHQGNSPDLEVYNVIRKQSEADLAETRP 300

Qy        301 DLKNISFRVCSGEATPDDMSCDYDNMAVNPSESGFVTLVSVESGFVTNDIYEFSPDQMGR 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 DLKNISFRVCSGEATPDDMSCDYDNMAVNPSESGFVTLVSVESGFVTNDIYEFSPDQMGR 360

Qy        361 SKESGWVENEIYGY 374
              ||||||||||||||
Db        361 SKESGWVENEIYGY 374
```

RESULT 2
US-10-149-819-15
; Sequence 15, Application US/10149819
; Publication No. US20030044913A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.
; APPLICANT: YUE, Henry
; APPLICANT: AZIMZAI, Yalda
; APPLICANT: TANG, Y. Tom
; APPLICANT: PATTERSON, Chandra
; APPLICANT: BAUGHN, Mariah R.
; APPLICANT: LU, Dyung Aina M.
; APPLICANT: SHAH, Purvi
; APPLICANT: LAL, Preeti
; APPLICANT: AU-YOUNG, Janice
; APPLICANT: BURFORD, Neil
; TITLE OF INVENTION: EXTRACELLULAR MATRIX AND CELL ADHESION MOLECULES
; FILE REFERENCE: PF-0760 PCT
; CURRENT APPLICATION NUMBER: US/10/149,819
; CURRENT FILING DATE: 2002-06-10
; PRIOR APPLICATION NUMBER: 60/172,852; 60/172,354
; PRIOR FILING DATE: 1999-12-10; 1999-12-16
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: PERL Program
; SEQ ID NO 15
; LENGTH: 374
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20030044913A1 3143411CD1
US-10-149-819-15

Query Match 100.0%; Score 2000; DB 14; Length 374;
Best Local Similarity 100.0%; Pred. No. 2.8e-179;
Matches 374; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MRPGTALQAVLLAVLLVGLRAATGRLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAK 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MRPGTALQAVLLAVLLVGLRAATGRLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAK 60

Qy         61 EACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 EACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWT 120

Qy        121 DGSISQFRNWYVDEPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSD 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 DGSISQFRNWYVDEPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSD 180

Qy        181 EKPAVPSREAEGEETELTTPVLPEETQEEDAKKTFKESREAALNLAYILIPSIPLLLLLV 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 EKPAVPSREAEGEETELTTPVLPEETQEEDAKKTFKESREAALNLAYILIPSIPLLLLLV 240

Qy        241 VTTVVCWVWICRKRKREQPDPSTKKQHTIWPSPHQGNSPDLEVYNVIRKQSEADLAETRP 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 VTTVVCWVWICRKRKREQPDPSTKKQHTIWPSPHQGNSPDLEVYNVIRKQSEADLAETRP 300

Qy        301 DLKNISFRVCSGEATPDDMSCDYDNMAVNPSESGFVTLVSVESGFVTNDIYEFSPDQMGR 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 DLKNISFRVCSGEATPDDMSCDYDNMAVNPSESGFVTLVSVESGFVTNDIYEFSPDQMGR 360

Qy        361 SKESGWVENEIYGY 374
              ||||||||||||||
Db        361 SKESGWVENEIYGY 374
```

RESULT 3
US-10-094-749-2090
; Sequence 2090, Application US/10094749
; Publication No. US20030219741A1
; GENERAL INFORMATION:
; APPLICANT: ISOGAI, TAKAO
; APPLICANT: SUGIYAMA, TOMOYASU
; APPLICANT: OTSUKI, TETSUJI
; APPLICANT: WAKAMATSU, AI
; APPLICANT: SATO, HIROYUKI
; APPLICANT: ISHII, SHIZUKO
; APPLICANT: YAMAMOTO, JUN-ICHI
; APPLICANT: ISONO, YUUKO
; APPLICANT: HIO, YURI
; APPLICANT: OTSUKA, KAORU
; APPLICANT: NAGAI, KEIICHI
; APPLICANT: IRIE, RYOTARO
; APPLICANT: TAMECHIKA, ICHIRO
; APPLICANT: SEKI, NAOHIKO
; APPLICANT: YOSHIKAWA, TSUTOMU
; APPLICANT: OTSUKA, MOTOYUKI
; APPLICANT: NAGAHARI, KENJI
; APPLICANT: MASUHO, YASUHIKO
; TITLE OF INVENTION: NOVEL FULL-LENGTH cDNA
; FILE REFERENCE: 084335/0160
; CURRENT APPLICATION NUMBER: US/10/094,749
; CURRENT FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 60/350,435
; PRIOR FILING DATE: 2002-01-24
; PRIOR APPLICATION NUMBER: JP 2001-328381
; PRIOR FILING DATE: 2001-09-14
; NUMBER OF SEQ ID NOS: 3381
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2090
; LENGTH: 374
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-094-749-2090

Query Match 100.0%; Score 2000; DB 15; Length 374;
Best Local Similarity 100.0%; Pred. No. 2.8e-179;
Matches 374; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MRPGTALQAVLLAVLLVGLRAATGRLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAK 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MRPGTALQAVLLAVLLVGLRAATGRLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAK 60

Qy         61 EACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 EACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWT 120

Qy        121 DGSISQFRNWYVDEPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSD 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 DGSISQFRNWYVDEPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSD 180

Qy        181 EKPAVPSREAEGEETELTTPVLPEETQEEDAKKTFKESREAALNLAYILIPSIPLLLLLV 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 EKPAVPSREAEGEETELTTPVLPEETQEEDAKKTFKESREAALNLAYILIPSIPLLLLLV 240

Qy        241 VTTVVCWVWICRKRKREQPDPSTKKQHTIWPSPHQGNSPDLEVYNVIRKQSEADLAETRP 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 VTTVVCWVWICRKRKREQPDPSTKKQHTIWPSPHQGNSPDLEVYNVIRKQSEADLAETRP 300

Qy        301 DLKNISFRVCSGEATPDDMSCDYDNMAVNPSESGFVTLVSVESGFVTNDIYEFSPDQMGR 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 DLKNISFRVCSGEATPDDMSCDYDNMAVNPSESGFVTLVSVESGFVTNDIYEFSPDQMGR 360

Qy        361 SKESGWVENEIYGY 374
              ||||||||||||||
Db        361 SKESGWVENEIYGY 374
```

```
Qy        361 SKESGWVENEIYGY 374
              ||||||||||||||
Db        361 SKESGWVENEIYGY 374


RESULT 4
US-10-264-237-2579
; Sequence 2579, Application US/10264237
; Publication No. US20040009491A1
; GENERAL INFORMATION:
;  APPLICANT: Birse et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
;  FILE REFERENCE: PA131P1
;  CURRENT APPLICATION NUMBER: US/10/264,237
;  CURRENT FILING DATE:  2002-10-04
;  PRIOR APPLICATION NUMBER: PCT/US01/16450
;  PRIOR FILING DATE: 2001-05-18
;  PRIOR APPLICATION NUMBER: US 60/205,515
;  PRIOR FILING DATE: 2000-05-19
;  NUMBER OF SEQ ID NOS: 2876
;  SOFTWARE: PatentIn Ver. 3.1
; SEQ ID NO 2579
;   LENGTH: 374
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-264-237-2579

  Query Match             100.0%;  Score 2000;  DB 15;  Length 374;
  Best Local Similarity   100.0%;  Pred. No. 2.8e-179;
  Matches  374;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MRPGTALQAVLLAVLLVGLRAATGRLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAK 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MRPGTALQAVLLAVLLVGLRAATGRLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAK 60

Qy         61 EACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 EACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWT 120

Qy        121 DGSISQFRNWYVDEPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSD 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 DGSISQFRNWYVDEPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSD 180

Qy        181 EKPAVPSREAEGEETELTTPVLPEETQEEDAKKTFKESREAALNLAYILIPSIPLLLLLV 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 EKPAVPSREAEGEETELTTPVLPEETQEEDAKKTFKESREAALNLAYILIPSIPLLLLLV 240

Qy        241 VTTVVCWVWICRKRKREQPDPSTKKQHTIWPSPHQGNSPDLEVYNVIRKQSEADLAETRP 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 VTTVVCWVWICRKRKREQPDPSTKKQHTIWPSPHQGNSPDLEVYNVIRKQSEADLAETRP 300

Qy        301 DLKNISFRVCSGEATPDDMSCDYDNMAVNPSESGFVTLVSVESGFVTNDIYEFSPDQMGR 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 DLKNISFRVCSGEATPDDMSCDYDNMAVNPSESGFVTLVSVESGFVTNDIYEFSPDQMGR 360

Qy        361 SKESGWVENEIYGY 374
              ||||||||||||||
Db        361 SKESGWVENEIYGY 374


RESULT 5
US-10-648-593-148
; Sequence 148, Application US/10648593
; Publication No. US20040106132A1
; GENERAL INFORMATION:
;  APPLICANT: Bristol-Myers Squibb Company
;  TITLE OF INVENTION: IDENTIFICATION OF GENES FOR PREDICTING ACTIVITY OF COMPOUNDS THAT
;  TITLE OF INVENTION:  INTERACT WITH AND/OR MODULATE PROTEIN TYROSINE KINASES AND/OR
;  TITLE OF INVENTION:  PROTEIN TYROSINE KINASE PATHWAYS IN BREAST CELLS
;  FILE REFERENCE: D0273 NP
;  CURRENT APPLICATION NUMBER: US/10/648,593
;  CURRENT FILING DATE:  2003-08-26
;  PRIOR APPLICATION NUMBER: 60/406,385
;  PRIOR FILING DATE: 2002-08-27
;  NUMBER OF SEQ ID NOS: 557
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 148
;   LENGTH: 374
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-648-593-148

  Query Match             100.0%;  Score 2000;  DB 16;  Length 374;
  Best Local Similarity   100.0%;  Pred. No. 2.8e-179;
  Matches  374;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MRPGTALQAVLLAVLLVGLRAATGRLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAK 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MRPGTALQAVLLAVLLVGLRAATGRLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAK 60

Qy         61 EACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 EACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWT 120

Qy        121 DGSISQFRNWYVDEPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSD 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 DGSISQFRNWYVDEPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSD 180

Qy        181 EKPAVPSREAEGEETELTTPVLPEETQEEDAKKTFKESREAALNLAYILIPSIPLLLLLV 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 EKPAVPSREAEGEETELTTPVLPEETQEEDAKKTFKESREAALNLAYILIPSIPLLLLLV 240

Qy        241 VTTVVCWVWICRKRKREQPDPSTKKQHTIWPSPHQGNSPDLEVYNVIRKQSEADLAETRP 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 VTTVVCWVWICRKRKREQPDPSTKKQHTIWPSPHQGNSPDLEVYNVIRKQSEADLAETRP 300

Qy        301 DLKNISFRVCSGEATPDDMSCDYDNMAVNPSESGFVTLVSVESGFVTNDIYEFSPDQMGR 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 DLKNISFRVCSGEATPDDMSCDYDNMAVNPSESGFVTLVSVESGFVTNDIYEFSPDQMGR 360

Qy        361 SKESGWVENEIYGY 374
              ||||||||||||||
Db        361 SKESGWVENEIYGY 374


RESULT 6
US-10-094-749-2142
; Sequence 2142, Application US/10094749
; Publication No. US20030219741A1
; GENERAL INFORMATION:
;  APPLICANT: ISOGAI, TAKAO
;  APPLICANT:  SUGIYAMA, TOMOYASU
;  APPLICANT:  OTSUKI, TETSUJI
;  APPLICANT:  WAKAMATSU, AI
;  APPLICANT:  SATO, HIROYUKI
;  APPLICANT:  ISHII, SHIZUKO
;  APPLICANT:  YAMAMOTO, JUN-ICHI
;  APPLICANT:  ISONO, YUUKO
;  APPLICANT:  HIO, YURI
;  APPLICANT:  OTSUKA, KAORU
;  APPLICANT:  NAGAI, KEIICHI
;  APPLICANT:  IRIE, RYOTARO
;  APPLICANT:  TAMECHIKA, ICHIRO
;  APPLICANT:  SEKI, NAOHIKO
;  APPLICANT:  YOSHIKAWA, TSUTOMU
;  APPLICANT:  OTSUKA, MOTOYUKI
;  APPLICANT:  NAGAHARI, KENJI
;  APPLICANT:  MASUHO, YASUHIKO
;  TITLE OF INVENTION: NOVEL FULL-LENGTH cDNA
;  FILE REFERENCE: 084335/0160
;  CURRENT APPLICATION NUMBER: US/10/094,749
;  CURRENT FILING DATE:  2002-03-12
;  PRIOR APPLICATION NUMBER: 60/350,435
;  PRIOR FILING DATE: 2002-01-24
;  PRIOR APPLICATION NUMBER: JP 2001-328381
```

```
;  PRIOR FILING DATE: 2001-09-14
;  NUMBER OF SEQ ID NOS: 3381
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2142
;    LENGTH: 374
;    TYPE: PRT
;    ORGANISM: Homo sapiens
US-10-094-749-2142

  Query Match             99.8%;  Score 1996;  DB 15;  Length 374;
  Best Local Similarity   99.7%;  Pred. No. 6.8e-179;
  Matches  373;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 MRPGTALQAVLLAVLLVGLRAATGRLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAK 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MRPGTALQAVLLAVLLVGLRAATGRLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAK 60

Qy         61 EACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 EACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWT 120

Qy        121 DGSISQFRNWYVDEPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSD 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 DGSISQFRNWYVDEPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSD 180

Qy        181 EKPAVPSREAEGEETELTTPVLPEETQEEDAKKTFKESREAALNLAYILIPSIPLLLLLV 240
              |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db        181 EKPAVPSREAEGEETELTTPVLPEETQEEDTKKTFKESREAALNLAYILIPSIPLLLLLV 240

Qy        241 VTTVVCWVWICRKRKREQPDPSTKKQHTIWPSPHQGNSPDLEVYNVIRKQSEADLAETRP 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 VTTVVCWVWICRKRKREQPDPSTKKQHTIWPSPHQGNSPDLEVYNVIRKQSEADLAETRP 300

Qy        301 DLKNISFRVCSGEATPDDMSCDYDNMAVNPSESGFVTLVSVESGFVTNDIYEFSPDQMGR 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 DLKNISFRVCSGEATPDDMSCDYDNMAVNPSESGFVTLVSVESGFVTNDIYEFSPDQMGR 360

Qy        361 SKESGWVENEIYGY 374
              ||||||||||||||
Db        361 SKESGWVENEIYGY 374


RESULT 7
US-10-351-334-166
; Sequence 166, Application US/10351334
; Publication No. US20040034196A1
; GENERAL INFORMATION:
;  APPLICANT: Komatsoulis et al
;  TITLE OF INVENTION: 98 Human Secreted Proteins
;  FILE REFERENCE: PZ031P2
;  CURRENT APPLICATION NUMBER: US/10/351,334
;  CURRENT FILING DATE:  2003-01-27
;  PRIOR APPLICATION NUMBER: 60/350,898
;  PRIOR FILING DATE: 2002-01-25
;  PRIOR APPLICATION NUMBER: 09/489,847
;  PRIOR FILING DATE: 2000-01-24
;  PRIOR APPLICATION NUMBER: PCT/US99/17130
;  PRIOR FILING DATE: 1999-07-29
;  PRIOR APPLICATION NUMBER: 60/094,657
;  PRIOR FILING DATE: 1998-07-30
;  PRIOR APPLICATION NUMBER: 60/095,486
;  PRIOR FILING DATE: 1998-08-05
;  PRIOR APPLICATION NUMBER: 60/096,319
;  PRIOR FILING DATE: 1998-08-12
;  PRIOR APPLICATION NUMBER: 60/095,454
;  PRIOR FILING DATE: 1998-08-06
;  PRIOR APPLICATION NUMBER: 60/095,455
;  PRIOR FILING DATE: 1998-08-06
;  NUMBER OF SEQ ID NOS: 376
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 166
;    LENGTH: 374
;    TYPE: PRT
;    ORGANISM: Homo sapiens
;    FEATURE:
;    NAME/KEY: SITE
;    LOCATION: (84)
;    OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
;    FEATURE:
;    NAME/KEY: SITE
;    LOCATION: (112)
;    OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-351-334-166

  Query Match             99.5%;  Score 1990;  DB 12;  Length 374;
  Best Local Similarity   99.5%;  Pred. No. 2.5e-178;
  Matches  372;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 MRPGTALQAVLLAVLLVGLRAATGRLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAK 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MRPGTALQAVLLAVLLVGLRAATGRLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAK 60

Qy         61 EACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWT 120
              |||||||||||||||||||| |||||||||||||||||||||||||| ||||||||
Db         61 EACRRDGGQLVSIESEDEQKLIEXFIENLLPSDGDFWIGLRRREEKQSNSTXCQDLYAWT 120

Qy        121 DGSISQFRNWYVDEPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSD 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 DGSISQFRNWYVDEPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSD 180

Qy        181 EKPAVPSREAEGEETELTTPVLPEETQEEDAKKTFKESREAALNLAYILIPSIPLLLLLV 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 EKPAVPSREAEGEETELTTPVLPEETQEEDAKKTFKESREAALNLAYILIPSIPLLLLLV 240

Qy        241 VTTVVCWVWICRKRKREQPDPSTKKQHTIWPSPHQGNSPDLEVYNVIRKQSEADLAETRP 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 VTTVVCWVWICRKRKREQPDPSTKKQHTIWPSPHQGNSPDLEVYNVIRKQSEADLAETRP 300

Qy        301 DLKNISFRVCSGEATPDDMSCDYDNMAVNPSESGFVTLVSVESGFVTNDIYEFSPDQMGR 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 DLKNISFRVCSGEATPDDMSCDYDNMAVNPSESGFVTLVSVESGFVTNDIYEFSPDQMGR 360

Qy        361 SKESGWVENEIYGY 374
              ||||||||||||||
Db        361 SKESGWVENEIYGY 374


RESULT 8
US-09-909-320-137
; Sequence 137, Application US/09909320
; Patent No. US20020132240A1
; GENERAL INFORMATION:
;  APPLICANT: Genentech, Inc.
;  APPLICANT:  Ashkenazi, Avi
;  APPLICANT:  Botstein, David
;  APPLICANT:  Desnoyers, Luc
;  APPLICANT:  Eaton, Dan L.
;  APPLICANT:  Ferrara, Napoleone
;  APPLICANT:  Filvaroff, Ellen
;  APPLICANT:  Fong, Sherman
;  APPLICANT:  Gao, Wei-Qiang
;  APPLICANT:  Gerber, Hanspeter
;  APPLICANT:  Gerritsen, Mary E.
;  APPLICANT:  Goddard, A.
;  APPLICANT:  Godowski, Paul J.
;  APPLICANT:  Grimaldi, Christopher J.
;  APPLICANT:  Gurney, Austin L.
;  APPLICANT:  Hillan, Kenneth, J.
;  APPLICANT:  Kljavin, Ivar J.
;  APPLICANT:  Mather, Jennie P.
;  APPLICANT:  Pan, James
;  APPLICANT:  Paoni, Nicholas F.
;  APPLICANT:  Roy, Margaret Ann
;  APPLICANT:  Stewart, Timothy A.
;  APPLICANT:  Tumas, Daniel
;  APPLICANT:  Williams, P. Mickey
```

; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/909,320
; CURRENT FILING DATE: 2002-01-04
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 137
; LENGTH: 382
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-909-320-137

Query Match 99.3%; Score 1986; DB 9; Length 382;
Best Local Similarity 97.9%; Pred. No. 6.1e-178;
Matches 374; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

```
Qy          1 MRPGTALQAVLLAVLLVGLRAATGRLLS--------GQPVCRGGTQRPCYKVIYFHDTSR 52
              |||||||||||||||||||||||||||||        ||||||||||||||||||||||
Db          1 MRPGTALQAVLLAVLLVGLRAATGRLLSASDLDLRGGQPVCRGGTQRPCYKVIYFHDTSR 60

Qy         53 RLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTA 112
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 RLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTA 120

Qy        113 CQDLYAWTDGSISQFRNWYVDEPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKN 172
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CQDLYAWTDGSISQFRNWYVDEPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKN 180

Qy        173 NFICKYSDEKPAVPSREAEGEETELTTPVLPEETQEEDAKKTFKESREAALNLAYILIPS 232
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 NFICKYSDEKPAVPSREAEGEETELTTPVLPEETQEEDAKKTFKESREAALNLAYILIPS 240

Qy        233 IPLLLLLVVTTVVCWVWICRKRKREQPDPSTKKQHTIWPSPHQGNSPDLEVYNVIRKQSE 292
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 IPLLLLLVVTTVVCWVWICRKRKREQPDPSTKKQHTIWPSPHQGNSPDLEVYNVIRKQSE 300

Qy        293 ADLAETRPDLKNISFRVCSGEATPDDMSCDYDNMAVNPSESGFVTLVSVESGFVTNDIYE 352
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        301 ADLAETRPDLKNISFRVCSGEATPDDMSCDYDNMAVNPSESGFVTLVSVESGFVTNDIYE 360

Qy        353 FSPDQMGRSKESGWVENEIYGY 374
              ||||||||||||||||||||||
Db        361 FSPDQMGRSKESGWVENEIYGY 382
```

RESULT 9
US-09-909-088B-137
; Sequence 137, Application US/09909088B
; Patent No. US20020146709A1
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/909,088B
; CURRENT FILING DATE: 2001-07-18
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999

; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 137
; LENGTH: 382
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-909-088B-137

Query Match 99.3%; Score 1986; DB 9; Length 382;
Best Local Similarity 97.9%; Pred. No. 6.1e-178;
Matches 374; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

```
Qy          1 MRPGTALQAVLLAVLLVGLRAATGRLLS--------GQPVCRGGTQRPCYKVIYFHDTSR 52
              |||||||||||||||||||||||||||        ||||||||||||||||||||||||
Db          1 MRPGTALQAVLLAVLLVGLRAATGRLLSASDLDLRGGQPVCRGGTQRPCYKVIYFHDTSR 60

Qy         53 RLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTA 112
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 RLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTA 120

Qy        113 CQDLYAWTDGSISQFRNWYVDEPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKN 172
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CQDLYAWTDGSISQFRNWYVDEPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKN 180

Qy        173 NFICKYSDEKPAVPSREAEGEETELTTPVLPEETQEEDAKKTFKESREAALNLAYILIPS 232
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 NFICKYSDEKPAVPSREAEGEETELTTPVLPEETQEEDAKKTFKESREAALNLAYILIPS 240

Qy        233 IPLLLLLVVTTVVCWVWICRKRKREQPDPSTKKQHTIWPSPHQGNSPDLEVYNVIRKQSE 292
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 IPLLLLLVVTTVVCWVWICRKRKREQPDPSTKKQHTIWPSPHQGNSPDLEVYNVIRKQSE 300

Qy        293 ADLAETRPDLKNISFRVCSGEATPDDMSCDYDNMAVNPSESGFVTLVSVESGFVTNDIYE 352
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ADLAETRPDLKNISFRVCSGEATPDDMSCDYDNMAVNPSESGFVTLVSVESGFVTNDIYE 360

Qy        353 FSPDQMGRSKESGWVENEIYGY 374
              ||||||||||||||||||||||
Db        361 FSPDQMGRSKESGWVENEIYGY 382
```

RESULT 10
US-09-905-291A-137
; Sequence 137, Application US/09905291A
; Patent No. US20020160374A1
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/905,291A
; CURRENT FILING DATE: 2001-07-12
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 137
; LENGTH: 382
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-905-291A-137

Query Match 99.3%; Score 1986; DB 9; Length 382;
Best Local Similarity 97.9%; Pred. No. 6.1e-178;
Matches 374; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

```
Qy          1 MRPGTALQAVLLAVLLVGLRAATGRLLS--------GQPVCRGGTQRPCYKVIYFHDTSR 52
              |||||||||||||||||||||||||||        ||||||||||||||||||||||||
Db          1 MRPGTALQAVLLAVLLVGLRAATGRLLSASDLDLRGGQPVCRGGTQRPCYKVIYFHDTSR 60

Qy         53 RLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTA 112
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 RLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTA 120

Qy        113 CQDLYAWTDGSISQFRNWYVDEPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKN 172
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CQDLYAWTDGSISQFRNWYVDEPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKN 180

Qy        173 NFICKYSDEKPAVPSREAEGEETELTTPVLPEETQEEDAKKTFKESREAALNLAYILIPS 232
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 NFICKYSDEKPAVPSREAEGEETELTTPVLPEETQEEDAKKTFKESREAALNLAYILIPS 240

Qy        233 IPLLLLLVVTTVVCWVWICRKRKREQPDPSTKKQHTIWPSPHQGNSPDLEVYNVIRKQSE 292
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 IPLLLLLVVTTVVCWVWICRKRKREQPDPSTKKQHTIWPSPHQGNSPDLEVYNVIRKQSE 300

Qy        293 ADLAETRPDLKNISFRVCSGEATPDDMSCDYDNMAVNPSESGFVTLVSVESGFVTNDIYE 352
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ADLAETRPDLKNISFRVCSGEATPDDMSCDYDNMAVNPSESGFVTLVSVESGFVTNDIYE 360
```

```
Qy        353 FSPDQMGRSKESGWVENEIYGY 374
              ||||||||||||||||||||||
Db        361 FSPDQMGRSKESGWVENEIYGY 382

RESULT 11
US-09-902-853-137
; Sequence 137, Application US/09902853
; Publication No. US20020192659A1
; GENERAL INFORMATION:
;  APPLICANT: Genentech, Inc.
;  APPLICANT:  Ashkenazi, Avi
;  APPLICANT:  Botstein, David
;  APPLICANT:  Desnoyers, Luc
;  APPLICANT:  Eaton, Dan L.
;  APPLICANT:  Ferrara, Napoleone
;  APPLICANT:  Filvaroff, Ellen
;  APPLICANT:  Fong, Sherman
;  APPLICANT:  Gao, Wei-Qiang
;  APPLICANT:  Gerber, Hanspeter
;  APPLICANT:  Gerritsen, Mary E.
;  APPLICANT:  Goddard, A.
;  APPLICANT:  Godowski, Paul J.
;  APPLICANT:  Grimaldi, Christopher J.
;  APPLICANT:  Gurney, Austin L.
;  APPLICANT:  Hillan, Kenneth, J.
;  APPLICANT:  Kljavin, Ivar J.
;  APPLICANT:  Mather, Jennie P.
;  APPLICANT:  Pan, James
;  APPLICANT:  Paoni, Nicholas F.
;  APPLICANT:  Roy, Margaret Ann
;  APPLICANT:  Stewart, Timothy A.
;  APPLICANT:  Tumas, Daniel
;  APPLICANT:  Williams, P. Mickey
;  APPLICANT:  Wood, William, I.
;  TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
;  TITLE OF INVENTION:  Acids Encoding the Same
;  FILE REFERENCE: 10466-14
;  CURRENT APPLICATION NUMBER: US/09/902,853
;  CURRENT FILING DATE:  2001-07-10
;  PRIOR APPLICATION NUMBER: US/09/665,350
;  PRIOR FILING DATE: 2000-09-18
;  PRIOR APPLICATION NUMBER: US 60/143,048
;  PRIOR FILING DATE: 1999-07-07
;  PRIOR APPLICATION NUMBER: US 60/145,698
;  PRIOR FILING DATE: 1999-07-26
;  PRIOR APPLICATION NUMBER: US 60/146,222
;  PRIOR FILING DATE: 1999-07-28
;  PRIOR APPLICATION NUMBER: PCT/US99/20594
;  PRIOR FILING DATE: 1999-09-08
;  PRIOR APPLICATION NUMBER: PCT/US99/20944
;  PRIOR FILING DATE: 1999-09-13
;  PRIOR APPLICATION NUMBER: PCT/US99/21090
;  PRIOR FILING DATE: 1999-09-15
;  PRIOR APPLICATION NUMBER: PCT/US99/21547
;  PRIOR FILING DATE: 1999-09-15
;  PRIOR APPLICATION NUMBER: PCT/US99/23089
;  PRIOR FILING DATE: 1999-10-05
;  PRIOR APPLICATION NUMBER: PCT/US99/28214
;  PRIOR FILING DATE: 1999-11-29
;  PRIOR APPLICATION NUMBER: PCT/US99/28313
;  PRIOR FILING DATE: 1999-11-30
;  PRIOR APPLICATION NUMBER: PCT/US99/28564
;  PRIOR FILING DATE: 1999-12-02
;  PRIOR APPLICATION NUMBER: PCT/US99/28565
;  PRIOR FILING DATE: 1999-12-02
;  PRIOR APPLICATION NUMBER: PCT/US99/30095
;  PRIOR FILING DATE: 1999-12-16
;  PRIOR APPLICATION NUMBER: PCT/US99/30911
;  PRIOR FILING DATE: 1999-12-20
;  PRIOR APPLICATION NUMBER: PCT/US99/30999
;  PRIOR FILING DATE: 1999-12-20
;  PRIOR APPLICATION NUMBER: PCT/US00/00219
;  PRIOR FILING DATE: 2000-01-05
;  NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 137
;   LENGTH: 382
;   TYPE: PRT
;   ORGANISM: Homo Sapien
US-09-902-853-137

  Query Match             99.3%;  Score 1986;  DB 9;  Length 382;
  Best Local Similarity   97.9%;  Pred. No. 6.1e-178;
  Matches  374;  Conservative    0;  Mismatches    0;  Indels    8;  Gaps    1;

Qy          1 MRPGTALQAVLLAVLLVGLRAATGRLLS--------GQPVCRGGTQRPCYKVIYFHDTSR 52
              |||||||||||||||||||||||||||        ||||||||||||||||||||||||
Db          1 MRPGTALQAVLLAVLLVGLRAATGRLLSASDLDLRGGQPVCRGGTQRPCYKVIYFHDTSR 60

Qy         53 RLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTA 112
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 RLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTA 120

Qy        113 CQDLYAWTDGSISQFRNWYVDEPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKN 172
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CQDLYAWTDGSISQFRNWYVDEPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKN 180

Qy        173 NFICKYSDEKPAVPSREAEGEETELTTPVLPEETQEEDAKKTFKESREAALNLAYILIPS 232
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 NFICKYSDEKPAVPSREAEGEETELTTPVLPEETQEEDAKKTFKESREAALNLAYILIPS 240

Qy        233 IPLLLLLVVTTVVCWVWICRKRKREQPDPSTKKQHTIWPSPHQGNSPDLEVYNVIRKQSE 292
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 IPLLLLLVVTTVVCWVWICRKRKREQPDPSTKKQHTIWPSPHQGNSPDLEVYNVIRKQSE 300

Qy        293 ADLAETRPDLKNISFRVCSGEATPDDMSCDYDNMAVNPSESGFVTLVSVESGFVTNDIYE 352
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ADLAETRPDLKNISFRVCSGEATPDDMSCDYDNMAVNPSESGFVTLVSVESGFVTNDIYE 360

Qy        353 FSPDQMGRSKESGWVENEIYGY 374
              ||||||||||||||||||||||
Db        361 FSPDQMGRSKESGWVENEIYGY 382

RESULT 12
US-09-907-824-137
; Sequence 137, Application US/09907824
; Publication No. US20020197671A1
; GENERAL INFORMATION:
;  APPLICANT: Genentech, Inc.
;  APPLICANT:  Ashkenazi, Avi
;  APPLICANT:  Botstein, David
;  APPLICANT:  Desnoyers, Luc
;  APPLICANT:  Eaton, Dan L.
;  APPLICANT:  Ferrara, Napoleone
;  APPLICANT:  Filvaroff, Ellen
;  APPLICANT:  Fong, Sherman
;  APPLICANT:  Gao, Wei-Qiang
;  APPLICANT:  Gerber, Hanspeter
;  APPLICANT:  Gerritsen, Mary E.
;  APPLICANT:  Goddard, A.
;  APPLICANT:  Godowski, Paul J.
;  APPLICANT:  Grimaldi, Christopher J.
;  APPLICANT:  Gurney, Austin L.
;  APPLICANT:  Hillan, Kenneth, J.
;  APPLICANT:  Kljavin, Ivar J.
;  APPLICANT:  Mather, Jennie P.
;  APPLICANT:  Pan, James
;  APPLICANT:  Paoni, Nicholas F.
;  APPLICANT:  Roy, Margaret Ann
;  APPLICANT:  Stewart, Timothy A.
;  APPLICANT:  Tumas, Daniel
;  APPLICANT:  Williams, P. Mickey
;  APPLICANT:  Wood, William, I.
;  TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
```

```
;  TITLE OF INVENTION:  Acids Encoding the Same
;  FILE REFERENCE: 10466-14
;  CURRENT APPLICATION NUMBER: US/09/907,824
;  CURRENT FILING DATE:  2001-07-17
;  PRIOR APPLICATION NUMBER: 09/665,350
;  PRIOR FILING DATE: 2000-09-18
;  PRIOR APPLICATION NUMBER: PCT/US00/04414
;  PRIOR FILING DATE: 2000-02-22
;  PRIOR APPLICATION NUMBER: US 60/143,048
;  PRIOR FILING DATE: 1999-07-07
;  PRIOR APPLICATION NUMBER: US 60/145,698
;  PRIOR FILING DATE: 1999-07-26
;  PRIOR APPLICATION NUMBER: US 60/146,222
;  PRIOR FILING DATE: 1999-07-28
;  PRIOR APPLICATION NUMBER: PCT/US99/20594
;  PRIOR FILING DATE: 1999-09-08
;  PRIOR APPLICATION NUMBER: PCT/US99/20944
;  PRIOR FILING DATE: 1999-09-13
;  PRIOR APPLICATION NUMBER: PCT/US99/21090
;  PRIOR FILING DATE: 1999-09-15
;  PRIOR APPLICATION NUMBER: PCT/US99/21547
;  PRIOR FILING DATE: 1999-09-15
;  PRIOR APPLICATION NUMBER: PCT/US99/23089
;  PRIOR FILING DATE: 1999-10-05
;  PRIOR APPLICATION NUMBER: PCT/US99/28214
;  PRIOR FILING DATE: 1999-11-29
;  PRIOR APPLICATION NUMBER: PCT/US99/28313
;  PRIOR FILING DATE: 1999-11-30
;  PRIOR APPLICATION NUMBER: PCT/US99/28564
;  PRIOR FILING DATE: 1999-12-02
;  PRIOR APPLICATION NUMBER: PCT/US99/28565
;  PRIOR FILING DATE: 1999-12-02
;  PRIOR APPLICATION NUMBER: PCT/US99/30095
;  PRIOR FILING DATE: 1999-12-16
;  PRIOR APPLICATION NUMBER: PCT/US99/30911
;  PRIOR FILING DATE: 1999-12-20
;  PRIOR APPLICATION NUMBER: PCT/US99/30999
;  PRIOR FILING DATE: 1999-12-20
;  PRIOR APPLICATION NUMBER: PCT/US00/00219
;  PRIOR FILING DATE: 2000-01-05
;  NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 137
;   LENGTH: 382
;   TYPE: PRT
;   ORGANISM: Homo Sapien
US-09-907-824-137

  Query Match             99.3%;  Score 1986;  DB 9;  Length 382;
  Best Local Similarity   97.9%;  Pred. No. 6.1e-178;
  Matches  374;  Conservative    0;  Mismatches    0;  Indels    8;  Gaps    1;

Qy          1 MRPGTALQAVLLAVLLVGLRAATGRLLS--------GQPVCRGGTQRPCYKVIYFHDTSR 52
              |||||||||||||||||||||||||||        ||||||||||||||||||||||||
Db          1 MRPGTALQAVLLAVLLVGLRAATGRLLSASDLDLRGGQPVCRGGTQRPCYKVIYFHDTSR 60

Qy         53 RLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTA 112
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 RLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTA 120

Qy        113 CQDLYAWTDGSISQFRNWYVDEPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKN 172
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CQDLYAWTDGSISQFRNWYVDEPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKN 180

Qy        173 NFICKYSDEKPAVPSREAEGEETELTTPVLPEETQEEDAKKTFKESREAALNLAYILIPS 232
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 NFICKYSDEKPAVPSREAEGEETELTTPVLPEETQEEDAKKTFKESREAALNLAYILIPS 240

Qy        233 IPLLLLLVVTTVVCWVWICRKRKREQPDPSTKKQHTIWPSPHQGNSPDLEVYNVIRKQSE 292
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 IPLLLLLVVTTVVCWVWICRKRKREQPDPSTKKQHTIWPSPHQGNSPDLEVYNVIRKQSE 300

Qy        293 ADLAETRPDLKNISFRVCSGEATPDDMSCDYDNMAVNPSESGFVTLVSVESGFVTNDIYE 352
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        301 ADLAETRPDLKNISFRVCSGEATPDDMSCDYDNMAVNPSESGFVTLVSVESGFVTNDIYE 360

Qy        353 FSPDQMGRSKESGWVENEIYGY 374
              ||||||||||||||||||||||
Db        361 FSPDQMGRSKESGWVENEIYGY 382


RESULT 13
US-09-907-841-137
; Sequence 137, Application US/09907841
; Publication No. US20020198366A1
; GENERAL INFORMATION:
;  APPLICANT: Genentech, Inc.
;  APPLICANT:  Ashkenazi, Avi
;  APPLICANT:  Botstein, David
;  APPLICANT:  Desnoyers, Luc
;  APPLICANT:  Eaton, Dan L.
;  APPLICANT:  Ferrara, Napoleone
;  APPLICANT:  Filvaroff, Ellen
;  APPLICANT:  Fong, Sherman
;  APPLICANT:  Gao, Wei-Qiang
;  APPLICANT:  Gerber, Hanspeter
;  APPLICANT:  Gerritsen, Mary E.
;  APPLICANT:  Goddard, A.
;  APPLICANT:  Godowski, Paul J.
;  APPLICANT:  Grimaldi, Christopher J.
;  APPLICANT:  Gurney, Austin L.
;  APPLICANT:  Hillan, Kenneth, J.
;  APPLICANT:  Kljavin, Ivar J.
;  APPLICANT:  Mather, Jennie P.
;  APPLICANT:  Pan, James
;  APPLICANT:  Paoni, Nicholas F.
;  APPLICANT:  Roy, Margaret Ann
;  APPLICANT:  Stewart, Timothy A.
;  APPLICANT:  Tumas, Daniel
;  APPLICANT:  Williams, P. Mickey
;  APPLICANT:  Wood, William, I.
;  TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
;  TITLE OF INVENTION:  Acids Encoding the Same
;  FILE REFERENCE: 10466-14
;  CURRENT APPLICATION NUMBER: US/09/907,841
;  CURRENT FILING DATE:  2001-11-20
;  PRIOR APPLICATION NUMBER: PCT/US00/04414
;  PRIOR FILING DATE: 2000-02-22
;  PRIOR APPLICATION NUMBER: US 60/143,048
;  PRIOR FILING DATE: 1999-07-07
;  PRIOR APPLICATION NUMBER: US 60/145,698
;  PRIOR FILING DATE: 1999-07-26
;  PRIOR APPLICATION NUMBER: US 60/146,222
;  PRIOR FILING DATE: 1999-07-28
;  PRIOR APPLICATION NUMBER: PCT/US99/20594
;  PRIOR FILING DATE: 1999-09-08
;  PRIOR APPLICATION NUMBER: PCT/US99/20944
;  PRIOR FILING DATE: 1999-09-13
;  PRIOR APPLICATION NUMBER: PCT/US99/21090
;  PRIOR FILING DATE: 1999-09-15
;  PRIOR APPLICATION NUMBER: PCT/US99/21547
;  PRIOR FILING DATE: 1999-09-15
;  PRIOR APPLICATION NUMBER: PCT/US99/23089
;  PRIOR FILING DATE: 1999-10-05
;  PRIOR APPLICATION NUMBER: PCT/US99/28214
;  PRIOR FILING DATE: 1999-11-29
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 137
;   LENGTH: 382
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-907-841-137

  Query Match             99.3%;  Score 1986;  DB 9;  Length 382;
  Best Local Similarity   97.9%;  Pred. No. 6.1e-178;
  Matches  374;  Conservative    0;  Mismatches    0;  Indels    8;  Gaps    1;
```

```
Qy          1 MRPGTALQAVLLAVLLVGLRAATGRLLS--------GQPVCRGGTQRPCYKVIYFHDTSR 52
              ||||||||||||||||||||||||||||        ||||||||||||||||||||||||
Db          1 MRPGTALQAVLLAVLLVGLRAATGRLLSASDLDLRGGQPVCRGGTQRPCYKVIYFHDTSR 60

Qy         53 RLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTA 112
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 RLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTA 120

Qy        113 CQDLYAWTDGSISQFRNWYVDEPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKN 172
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CQDLYAWTDGSISQFRNWYVDEPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKN 180

Qy        173 NFICKYSDEKPAVPSREAEGEETELTTPVLPEETQEEDAKKTFKESREAALNLAYILIPS 232
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 NFICKYSDEKPAVPSREAEGEETELTTPVLPEETQEEDAKKTFKESREAALNLAYILIPS 240

Qy        233 IPLLLLLVVTTVVCWVWICRKRKREQPDPSTKKQHTIWPSPHQGNSPDLEVYNVIRKQSE 292
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 IPLLLLLVVTTVVCWVWICRKRKREQPDPSTKKQHTIWPSPHQGNSPDLEVYNVIRKQSE 300

Qy        293 ADLAETRPDLKNISFRVCSGEATPDDMSCDYDNMAVNPSESGFVTLVSVESGFVTNDIYE 352
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ADLAETRPDLKNISFRVCSGEATPDDMSCDYDNMAVNPSESGFVTLVSVESGFVTNDIYE 360

Qy        353 FSPDQMGRSKESGWVENEIYGY 374
              ||||||||||||||||||||||
Db        361 FSPDQMGRSKESGWVENEIYGY 382
```

RESULT 14
US-09-904-011-137
; Sequence 137, Application US/09904011
; Publication No. US20030003530A1
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/904,011
; CURRENT FILING DATE: 2001-07-11
; PRIOR APPLICATION NUMBER: 09/665,350
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 137
; LENGTH: 382
; TYPE: PRT
; ORGANISM: Homo Sapien
US-09-904-011-137

Query Match 99.3%; Score 1986; DB 10; Length 382;
Best Local Similarity 97.9%; Pred. No. 6.1e-178;
Matches 374; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

```
Qy          1 MRPGTALQAVLLAVLLVGLRAATGRLLS--------GQPVCRGGTQRPCYKVIYFHDTSR 52
              ||||||||||||||||||||||||||||        ||||||||||||||||||||||||
Db          1 MRPGTALQAVLLAVLLVGLRAATGRLLSASDLDLRGGQPVCRGGTQRPCYKVIYFHDTSR 60

Qy         53 RLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTA 112
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 RLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTA 120

Qy        113 CQDLYAWTDGSISQFRNWYVDEPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKN 172
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CQDLYAWTDGSISQFRNWYVDEPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKN 180

Qy        173 NFICKYSDEKPAVPSREAEGEETELTTPVLPEETQEEDAKKTFKESREAALNLAYILIPS 232
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 NFICKYSDEKPAVPSREAEGEETELTTPVLPEETQEEDAKKTFKESREAALNLAYILIPS 240

Qy        233 IPLLLLLVVTTVVCWVWICRKRKREQPDPSTKKQHTIWPSPHQGNSPDLEVYNVIRKQSE 292
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 IPLLLLLVVTTVVCWVWICRKRKREQPDPSTKKQHTIWPSPHQGNSPDLEVYNVIRKQSE 300

Qy        293 ADLAETRPDLKNISFRVCSGEATPDDMSCDYDNMAVNPSESGFVTLVSVESGFVTNDIYE 352
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ADLAETRPDLKNISFRVCSGEATPDDMSCDYDNMAVNPSESGFVTLVSVESGFVTNDIYE 360

Qy        353 FSPDQMGRSKESGWVENEIYGY 374
              ||||||||||||||||||||||
Db        361 FSPDQMGRSKESGWVENEIYGY 382
```

RESULT 15
US-09-906-742-137
; Sequence 137, Application US/09906742
; Publication No. US20030023054A1
; GENERAL INFORMATION:

; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/906,742
; CURRENT FILING DATE: 2001-07-16
; PRIOR APPLICATION NUMBER: 09/665,350
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 137
; LENGTH: 382
; TYPE: PRT
; ORGANISM: Homo Sapien
US-09-906-742-137

Query Match 99.3%; Score 1986; DB 10; Length 382;
Best Local Similarity 97.9%; Pred. No. 6.1e-178;
Matches 374; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

```
Qy          1 MRPGTALQAVLLAVLLVGLRAATGRLLS--------GQPVCRGGTQRPCYKVIYFHDTSR 52
              ||||||||||||||||||||||||||||        ||||||||||||||||||||||||
Db          1 MRPGTALQAVLLAVLLVGLRAATGRLLSASDLDLRGGQPVCRGGTQRPCYKVIYFHDTSR 60

Qy         53 RLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTA 112
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 RLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTA 120

Qy        113 CQDLYAWTDGSISQFRNWYVDEPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKN 172
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CQDLYAWTDGSISQFRNWYVDEPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKN 180

Qy        173 NFICKYSDEKPAVPSREAEGEETELTTPVLPEETQEEDAKKTFKESREAALNLAYILIPS 232
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 NFICKYSDEKPAVPSREAEGEETELTTPVLPEETQEEDAKKTFKESREAALNLAYILIPS 240

Qy        233 IPLLLLLVVTTVVCWVWICRKRKREQPDPSTKKQHTIWPSPHQGNSPDLEVYNVIRKQSE 292
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 IPLLLLLVVTTVVCWVWICRKRKREQPDPSTKKQHTIWPSPHQGNSPDLEVYNVIRKQSE 300

Qy        293 ADLAETRPDLKNISFRVCSGEATPDDMSCDYDNMAVNPSESGFVTLVSVESGFVTNDIYE 352
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ADLAETRPDLKNISFRVCSGEATPDDMSCDYDNMAVNPSESGFVTLVSVESGFVTNDIYE 360

Qy        353 FSPDQMGRSKESGWVENEIYGY 374
              ||||||||||||||||||||||
Db        361 FSPDQMGRSKESGWVENEIYGY 382
```

Search completed: September 9, 2004, 22:34:10
Job time : 133 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: September 9, 2004, 22:21:36 ; Search time 417 Seconds
(without alignments)
875.405 Million cell updates/sec

Title: US-09-887-855-2
Perfect score: 2000
Sequence: 1 MRPGTALQAVLLAVLLVGLR.....PDQMGRSKESGWVENEIYGY 374

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 6019581 seqs, 976053577 residues

Total number of hits satisfying chosen parameters: 6019581

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending_Patents_AA_Main:*
1: /cgn2_6/ptodata/2/paa/PCTUS_COMB.pep:*
2: /cgn2_6/ptodata/2/paa/US06_COMB.pep:*
3: /cgn2_6/ptodata/2/paa/US07_COMB.pep:*
4: /cgn2_6/ptodata/2/paa/US080_COMB.pep:*
5: /cgn2_6/ptodata/2/paa/US081_COMB.pep:*
6: /cgn2_6/ptodata/2/paa/US082_COMB.pep:*
7: /cgn2_6/ptodata/2/paa/US083_COMB.pep:*
8: /cgn2_6/ptodata/2/paa/US084_COMB.pep:*
9: /cgn2_6/ptodata/2/paa/US085_COMB.pep:*
10: /cgn2_6/ptodata/2/paa/US086_COMB.pep:*
11: /cgn2_6/ptodata/2/paa/US087_COMB.pep:*
12: /cgn2_6/ptodata/2/paa/US088_COMB.pep:*
13: /cgn2_6/ptodata/2/paa/US089_COMB.pep:*
14: /cgn2_6/ptodata/2/paa/US090_COMB.pep:*
15: /cgn2_6/ptodata/2/paa/US091_COMB.pep:*
16: /cgn2_6/ptodata/2/paa/US092_COMB.pep:*
17: /cgn2_6/ptodata/2/paa/US093_COMB.pep:*
18: /cgn2_6/ptodata/2/paa/US094_COMB.pep:*
19: /cgn2_6/ptodata/2/paa/US095_COMB.pep:*
20: /cgn2_6/ptodata/2/paa/US096_COMB.pep:*
21: /cgn2_6/ptodata/2/paa/US097A_COMB.pep:*
22: /cgn2_6/ptodata/2/paa/US097B_COMB.pep:*
23: /cgn2_6/ptodata/2/paa/US098_COMB.pep:*
24: /cgn2_6/ptodata/2/paa/US099A_COMB.pep:*
25: /cgn2_6/ptodata/2/paa/US099B_COMB.pep:*
26: /cgn2_6/ptodata/2/paa/US100_COMB.pep:*
27: /cgn2_6/ptodata/2/paa/US101_COMB.pep:*
28: /cgn2_6/ptodata/2/paa/US102_COMB.pep:*
29: /cgn2_6/ptodata/2/paa/US103_COMB.pep:*
30: /cgn2_6/ptodata/2/paa/US104_COMB.pep:*
31: /cgn2_6/ptodata/2/paa/US106_COMB.pep:*
32: /cgn2_6/ptodata/2/paa/US107_COMB.pep:*
33: /cgn2_6/ptodata/2/paa/US60_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2000	100.0	374	1	PCT-US00-32990-15	Sequence 15, Appl
2	2000	100.0	374	1	PCT-US01-16450-2579	Sequence 2579, Ap
3	2000	100.0	374	1	PCT-US01-16450A-2579	Sequence 2579, Ap
4	2000	100.0	374	1	PCT-US03-26491-148	Sequence 148, App
5	2000	100.0	374	23	US-09-887-855-2	Sequence 2, Appli
6	2000	100.0	374	26	US-10-094-749-2090	Sequence 2090, Ap
7	2000	100.0	374	27	US-10-149-819-15	Sequence 15, Appl
8	2000	100.0	374	28	US-10-264-237-2579	Sequence 2579, Ap
9	2000	100.0	374	31	US-10-648-593-148	Sequence 148, App
10	2000	100.0	374	33	US-60-172-354-15	Sequence 15, Appl
11	1996	99.8	374	26	US-10-094-749-2142	Sequence 2142, Ap
12	1990	99.5	374	1	PCT-US99-17130-163	Sequence 163, App
13	1990	99.5	374	29	US-10-351-334-166	Sequence 166, App
14	1986	99.3	382	17	US-09-380-139A-137	Sequence 137, App
15	1986	99.3	382	18	US-09-423-844-137	Sequence 137, App
16	1986	99.3	382	20	US-09-664-610B-137	Sequence 137, App
17	1986	99.3	382	20	US-09-665-350-137	Sequence 137, App
18	1986	99.3	382	20	US-09-665-350B-137	Sequence 137, App
19	1986	99.3	382	24	US-09-902-572A-137	Sequence 137, App
20	1986	99.3	382	24	US-09-902-615-137	Sequence 137, App
21	1986	99.3	382	24	US-09-902-634-137	Sequence 137, App
22	1986	99.3	382	24	US-09-902-634A-137	Sequence 137, App
23	1986	99.3	382	24	US-09-902-654-137	Sequence 137, App
24	1986	99.3	382	24	US-09-902-692-137	Sequence 137, App
25	1986	99.3	382	24	US-09-902-713-137	Sequence 137, App
26	1986	99.3	382	24	US-09-902-713B-137	Sequence 137, App
27	1986	99.3	382	24	US-09-902-736-137	Sequence 137, App
28	1986	99.3	382	24	US-09-902-736A-137	Sequence 137, App
29	1986	99.3	382	24	US-09-902-759-137	Sequence 137, App
30	1986	99.3	382	24	US-09-902-853-137	Sequence 137, App
31	1986	99.3	382	24	US-09-902-853A-137	Sequence 137, App
32	1986	99.3	382	24	US-09-902-903-137	Sequence 137, App
33	1986	99.3	382	24	US-09-902-979-137	Sequence 137, App
34	1986	99.3	382	24	US-09-902-979A-137	Sequence 137, App
35	1986	99.3	382	24	US-09-903-520-137	Sequence 137, App
36	1986	99.3	382	24	US-09-903-520A-137	Sequence 137, App
37	1986	99.3	382	24	US-09-903-562-137	Sequence 137, App
38	1986	99.3	382	24	US-09-903-562A-137	Sequence 137, App
39	1986	99.3	382	24	US-09-903-562B-137	Sequence 137, App
40	1986	99.3	382	24	US-09-903-603A-137	Sequence 137, App
41	1986	99.3	382	24	US-09-903-640-137	Sequence 137, App
42	1986	99.3	382	24	US-09-903-640A-137	Sequence 137, App
43	1986	99.3	382	24	US-09-903-663-137	Sequence 137, App
44	1986	99.3	382	24	US-09-903-663A-137	Sequence 137, App
45	1986	99.3	382	24	US-09-903-749A-137	Sequence 137, App

ALIGNMENTS

RESULT 1
PCT-US00-32990-15
; Sequence 15, Application PC/TUS0032990
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.
; APPLICANT: YUE, Henry
; APPLICANT: AZIMZAI, Yalda
; APPLICANT: TANG, Y. Tom
; APPLICANT: PATTERSON, Chandra
; APPLICANT: BAUGHN, Mariah R.
; APPLICANT: LU, Dyung Aina M.
; APPLICANT: SHAH, Purvi
; APPLICANT: LAL, Preeti
; APPLICANT: AU-YOUNG, Janice
; APPLICANT: BURFORD, Neil
; TITLE OF INVENTION: EXTRACELLULAR MATRIX AND CELL ADHESION MOLECULES
; FILE REFERENCE: PF-0760 PCT
; CURRENT APPLICATION NUMBER: PCT/US00/32990
; CURRENT FILING DATE: 2000-12-05
; PRIOR APPLICATION NUMBER: 60/172,852; 60/172,354
; PRIOR FILING DATE: 1999-12-10; 1999-12-16
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: PERL Program

```
; SEQ ID NO 15
;   LENGTH: 374
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No: 3143411CD1
PCT-US00-32990-15

  Query Match            100.0%;  Score 2000;  DB 1;  Length 374;
  Best Local Similarity  100.0%;  Pred. No. 3.4e-184;
  Matches  374;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MRPGTALQAVLLAVLLVGLRAATGRLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAK 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MRPGTALQAVLLAVLLVGLRAATGRLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAK 60

Qy         61 EACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 EACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWT 120

Qy        121 DGSISQFRNWYVDEPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSD 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 DGSISQFRNWYVDEPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSD 180

Qy        181 EKPAVPSREAEGEETELTTPVLPEETQEEDAKKTFKESREAALNLAYILIPSIPLLLLLV 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 EKPAVPSREAEGEETELTTPVLPEETQEEDAKKTFKESREAALNLAYILIPSIPLLLLLV 240

Qy        241 VTTVVCWVWICRKRKREQPDPSTKKQHTIWPSPHQGNSPDLEVYNVIRKQSEADLAETRP 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 VTTVVCWVWICRKRKREQPDPSTKKQHTIWPSPHQGNSPDLEVYNVIRKQSEADLAETRP 300

Qy        301 DLKNISFRVCSGEATPDDMSCDYDNMAVNPSESGFVTLVSVESGFVTNDIYEFSPDQMGR 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 DLKNISFRVCSGEATPDDMSCDYDNMAVNPSESGFVTLVSVESGFVTNDIYEFSPDQMGR 360

Qy        361 SKESGWVENEIYGY 374
              ||||||||||||||
Db        361 SKESGWVENEIYGY 374


RESULT 2
PCT-US01-16450-2579
; Sequence 2579, Application PC/TUS0116450
; GENERAL INFORMATION:
;  APPLICANT: Human Genome Sciences, Inc.
;  TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
;  FILE REFERENCE: PA131PCT
;  CURRENT APPLICATION NUMBER: PCT/US01/16450
;  CURRENT FILING DATE:  2001-06-01
;  PRIOR APPLICATION NUMBER: 60/205,515
;  PRIOR FILING DATE: 2000-05-19
;  NUMBER OF SEQ ID NOS: 2820
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2579
;   LENGTH: 374
;   TYPE: PRT
;   ORGANISM: Homo sapiens
PCT-US01-16450-2579

  Query Match            100.0%;  Score 2000;  DB 1;  Length 374;
  Best Local Similarity  100.0%;  Pred. No. 3.4e-184;
  Matches  374;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MRPGTALQAVLLAVLLVGLRAATGRLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAK 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MRPGTALQAVLLAVLLVGLRAATGRLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAK 60

Qy         61 EACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 EACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWT 120

Qy        121 DGSISQFRNWYVDEPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSD 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 DGSISQFRNWYVDEPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSD 180

Qy        181 EKPAVPSREAEGEETELTTPVLPEETQEEDAKKTFKESREAALNLAYILIPSIPLLLLLV 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 EKPAVPSREAEGEETELTTPVLPEETQEEDAKKTFKESREAALNLAYILIPSIPLLLLLV 240

Qy        241 VTTVVCWVWICRKRKREQPDPSTKKQHTIWPSPHQGNSPDLEVYNVIRKQSEADLAETRP 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 VTTVVCWVWICRKRKREQPDPSTKKQHTIWPSPHQGNSPDLEVYNVIRKQSEADLAETRP 300

Qy        301 DLKNISFRVCSGEATPDDMSCDYDNMAVNPSESGFVTLVSVESGFVTNDIYEFSPDQMGR 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 DLKNISFRVCSGEATPDDMSCDYDNMAVNPSESGFVTLVSVESGFVTNDIYEFSPDQMGR 360

Qy        361 SKESGWVENEIYGY 374
              ||||||||||||||
Db        361 SKESGWVENEIYGY 374


RESULT 3
PCT-US01-16450A-2579
; Sequence 2579, Application PC/TUS0116450A
; GENERAL INFORMATION:
;  APPLICANT: Human Genome Sciences, Inc.
;  TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
;  FILE REFERENCE: PA131PCT
;  CURRENT APPLICATION NUMBER: PCT/US01/16450A
;  CURRENT FILING DATE:  2001-05-18
;  PRIOR APPLICATION NUMBER: 60/205,515
;  PRIOR FILING DATE: 2000-05-19
;  NUMBER OF SEQ ID NOS: 2820
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2579
;   LENGTH: 374
;   TYPE: PRT
;   ORGANISM: Homo sapiens
PCT-US01-16450A-2579

  Query Match            100.0%;  Score 2000;  DB 1;  Length 374;
  Best Local Similarity  100.0%;  Pred. No. 3.4e-184;
  Matches  374;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MRPGTALQAVLLAVLLVGLRAATGRLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAK 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MRPGTALQAVLLAVLLVGLRAATGRLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAK 60

Qy         61 EACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 EACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWT 120

Qy        121 DGSISQFRNWYVDEPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSD 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 DGSISQFRNWYVDEPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSD 180

Qy        181 EKPAVPSREAEGEETELTTPVLPEETQEEDAKKTFKESREAALNLAYILIPSIPLLLLLV 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 EKPAVPSREAEGEETELTTPVLPEETQEEDAKKTFKESREAALNLAYILIPSIPLLLLLV 240

Qy        241 VTTVVCWVWICRKRKREQPDPSTKKQHTIWPSPHQGNSPDLEVYNVIRKQSEADLAETRP 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 VTTVVCWVWICRKRKREQPDPSTKKQHTIWPSPHQGNSPDLEVYNVIRKQSEADLAETRP 300

Qy        301 DLKNISFRVCSGEATPDDMSCDYDNMAVNPSESGFVTLVSVESGFVTNDIYEFSPDQMGR 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 DLKNISFRVCSGEATPDDMSCDYDNMAVNPSESGFVTLVSVESGFVTNDIYEFSPDQMGR 360

Qy        361 SKESGWVENEIYGY 374
              ||||||||||||||
Db        361 SKESGWVENEIYGY 374
```

```
RESULT 4
PCT-US03-26491-148
; Sequence 148, Application PC/TUS0326491
; GENERAL INFORMATION:
;  APPLICANT: Bristol-Myers Squibb Company
;  TITLE OF INVENTION: IDENTIFICATION OF GENES FOR PREDICTING ACTIVITY OF COMPOUNDS THAT
;  TITLE OF INVENTION:  INTERACT WITH AND/OR MODULATE PROTEIN TYROSINE KINASES AND/OR
;  TITLE OF INVENTION:  PROTEIN TYROSINE KINASE PATHWAYS IN BREAST CELLS
;  FILE REFERENCE: D0273 PCT
;  CURRENT APPLICATION NUMBER: PCT/US03/26491
;  CURRENT FILING DATE:  2003-08-26
;  PRIOR APPLICATION NUMBER: 60/406,385
;  PRIOR FILING DATE: 2002-08-27
;  NUMBER OF SEQ ID NOS: 557
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 148
;   LENGTH: 374
;   TYPE: PRT
;   ORGANISM: Homo sapiens
PCT-US03-26491-148

  Query Match             100.0%;  Score 2000;  DB 1;  Length 374;
  Best Local Similarity   100.0%;  Pred. No. 3.4e-184;
  Matches  374;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MRPGTALQAVLLAVLLVGLRAATGRLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAK 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MRPGTALQAVLLAVLLVGLRAATGRLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAK 60

Qy         61 EACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 EACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWT 120

Qy        121 DGSISQFRNWYVDEPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSD 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 DGSISQFRNWYVDEPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSD 180

Qy        181 EKPAVPSREAEGEETELTTPVLPEETQEEDAKKTFKESREAALNLAYILIPSIPLLLLLV 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 EKPAVPSREAEGEETELTTPVLPEETQEEDAKKTFKESREAALNLAYILIPSIPLLLLLV 240

Qy        241 VTTVVCWVWICRKRKREQPDPSTKKQHTIWPSPHQGNSPDLEVYNVIRKQSEADLAETRP 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 VTTVVCWVWICRKRKREQPDPSTKKQHTIWPSPHQGNSPDLEVYNVIRKQSEADLAETRP 300

Qy        301 DLKNISFRVCSGEATPDDMSCDYDNMAVNPSESGFVTLVSVESGFVTNDIYEFSPDQMGR 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 DLKNISFRVCSGEATPDDMSCDYDNMAVNPSESGFVTLVSVESGFVTNDIYEFSPDQMGR 360

Qy        361 SKESGWVENEIYGY 374
              ||||||||||||||
Db        361 SKESGWVENEIYGY 374


RESULT 5
US-09-887-855-2
; Sequence 2, Application US/09887855
; GENERAL INFORMATION:
;  APPLICANT: Immunex Corporation
;  APPLICANT:  Anderson, Dirk M
;  TITLE OF INVENTION: LECTIN SS3939 DNA AND POLYPEPTIDES
;  FILE REFERENCE: 2883-US
;  CURRENT APPLICATION NUMBER: US/09/887,855
;  CURRENT FILING DATE:  2001-06-22
;  NUMBER OF SEQ ID NOS: 9
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
;   LENGTH: 374
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-887-855-2

  Query Match             100.0%;  Score 2000;  DB 23;  Length 374;
  Best Local Similarity   100.0%;  Pred. No. 3.4e-184;
  Matches  374;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MRPGTALQAVLLAVLLVGLRAATGRLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAK 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MRPGTALQAVLLAVLLVGLRAATGRLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAK 60

Qy         61 EACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 EACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWT 120

Qy        121 DGSISQFRNWYVDEPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSD 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 DGSISQFRNWYVDEPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSD 180

Qy        181 EKPAVPSREAEGEETELTTPVLPEETQEEDAKKTFKESREAALNLAYILIPSIPLLLLLV 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 EKPAVPSREAEGEETELTTPVLPEETQEEDAKKTFKESREAALNLAYILIPSIPLLLLLV 240

Qy        241 VTTVVCWVWICRKRKREQPDPSTKKQHTIWPSPHQGNSPDLEVYNVIRKQSEADLAETRP 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 VTTVVCWVWICRKRKREQPDPSTKKQHTIWPSPHQGNSPDLEVYNVIRKQSEADLAETRP 300

Qy        301 DLKNISFRVCSGEATPDDMSCDYDNMAVNPSESGFVTLVSVESGFVTNDIYEFSPDQMGR 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 DLKNISFRVCSGEATPDDMSCDYDNMAVNPSESGFVTLVSVESGFVTNDIYEFSPDQMGR 360

Qy        361 SKESGWVENEIYGY 374
              ||||||||||||||
Db        361 SKESGWVENEIYGY 374


RESULT 6
US-10-094-749-2090
; Sequence 2090, Application US/10094749
; GENERAL INFORMATION:
;  APPLICANT: ISOGAI, TAKAO
;  APPLICANT:  SUGIYAMA, TOMOYASU
;  APPLICANT:  OTSUKI, TETSUJI
;  APPLICANT:  WAKAMATSU, AI
;  APPLICANT:  SATO, HIROYUKI
;  APPLICANT:  ISHII, SHIZUKO
;  APPLICANT:  YAMAMOTO, JUN-ICHI
;  APPLICANT:  ISONO, YUUKO
;  APPLICANT:  HIO, YURI
;  APPLICANT:  OTSUKA, KAORU
;  APPLICANT:  NAGAI, KEIICHI
;  APPLICANT:  IRIE, RYOTARO
;  APPLICANT:  TAMECHIKA, ICHIRO
;  APPLICANT:  SEKI, NAOHIKO
;  APPLICANT:  YOSHIKAWA, TSUTOMU
;  APPLICANT:  OTSUKA, MOTOYUKI
;  APPLICANT:  NAGAHARI, KENJI
;  APPLICANT:  MASUHO, YASUHIKO
;  TITLE OF INVENTION: NOVEL FULL-LENGTH cDNA
;  FILE REFERENCE: 084335/0160
;  CURRENT APPLICATION NUMBER: US/10/094,749
;  CURRENT FILING DATE:  2002-03-12
;  PRIOR APPLICATION NUMBER: 60/350,435
;  PRIOR FILING DATE: 2002-01-24
;  PRIOR APPLICATION NUMBER: JP 2001-328381
;  PRIOR FILING DATE: 2001-09-14
;  NUMBER OF SEQ ID NOS: 3381
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2090
;   LENGTH: 374
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-094-749-2090

  Query Match             100.0%;  Score 2000;  DB 26;  Length 374;
  Best Local Similarity   100.0%;  Pred. No. 3.4e-184;
  Matches  374;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy          1 MRPGTALQAVLLAVLLVGLRAATGRLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAK 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MRPGTALQAVLLAVLLVGLRAATGRLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAK 60

Qy         61 EACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 EACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWT 120

Qy        121 DGSISQFRNWYVDEPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSD 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 DGSISQFRNWYVDEPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSD 180

Qy        181 EKPAVPSREAEGEETELTTPVLPEETQEEDAKKTFKESREAALNLAYILIPSIPLLLLLV 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 EKPAVPSREAEGEETELTTPVLPEETQEEDAKKTFKESREAALNLAYILIPSIPLLLLLV 240

Qy        241 VTTVVCWVWICRKRKREQPDPSTKKQHTIWPSPHQGNSPDLEVYNVIRKQSEADLAETRP 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 VTTVVCWVWICRKRKREQPDPSTKKQHTIWPSPHQGNSPDLEVYNVIRKQSEADLAETRP 300

Qy        301 DLKNISFRVCSGEATPDDMSCDYDNMAVNPSESGFVTLVSVESGFVTNDIYEFSPDQMGR 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 DLKNISFRVCSGEATPDDMSCDYDNMAVNPSESGFVTLVSVESGFVTNDIYEFSPDQMGR 360

Qy        361 SKESGWVENEIYGY 374
              ||||||||||||||
Db        361 SKESGWVENEIYGY 374
```

RESULT 7
US-10-149-819-15
; Sequence 15, Application US/10149819
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.
; APPLICANT: YUE, Henry
; APPLICANT: AZIMZAI, Yalda
; APPLICANT: TANG, Y. Tom
; APPLICANT: PATTERSON, Chandra
; APPLICANT: BAUGHN, Mariah R.
; APPLICANT: LU, Dyung Aina M.
; APPLICANT: SHAH, Purvi
; APPLICANT: LAL, Preeti
; APPLICANT: AU-YOUNG, Janice
; APPLICANT: BURFORD, Neil
; TITLE OF INVENTION: EXTRACELLULAR MATRIX AND CELL ADHESION MOLECULES
; FILE REFERENCE: PF-0760 PCT
; CURRENT APPLICATION NUMBER: US/10/149,819
; CURRENT FILING DATE: 2002-06-10
; PRIOR APPLICATION NUMBER: 60/172,852; 60/172,354
; PRIOR FILING DATE: 1999-12-10; 1999-12-16
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: PERL Program
; SEQ ID NO 15
; LENGTH: 374
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No: 3143411CD1
US-10-149-819-15

Query Match 100.0%; Score 2000; DB 27; Length 374;
Best Local Similarity 100.0%; Pred. No. 3.4e-184;
Matches 374; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MRPGTALQAVLLAVLLVGLRAATGRLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAK 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MRPGTALQAVLLAVLLVGLRAATGRLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAK 60

Qy         61 EACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 EACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWT 120

Qy        121 DGSISQFRNWYVDEPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSD 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 DGSISQFRNWYVDEPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSD 180

Qy        181 EKPAVPSREAEGEETELTTPVLPEETQEEDAKKTFKESREAALNLAYILIPSIPLLLLLV 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 EKPAVPSREAEGEETELTTPVLPEETQEEDAKKTFKESREAALNLAYILIPSIPLLLLLV 240

Qy        241 VTTVVCWVWICRKRKREQPDPSTKKQHTIWPSPHQGNSPDLEVYNVIRKQSEADLAETRP 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 VTTVVCWVWICRKRKREQPDPSTKKQHTIWPSPHQGNSPDLEVYNVIRKQSEADLAETRP 300

Qy        301 DLKNISFRVCSGEATPDDMSCDYDNMAVNPSESGFVTLVSVESGFVTNDIYEFSPDQMGR 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 DLKNISFRVCSGEATPDDMSCDYDNMAVNPSESGFVTLVSVESGFVTNDIYEFSPDQMGR 360

Qy        361 SKESGWVENEIYGY 374
              ||||||||||||||
Db        361 SKESGWVENEIYGY 374
```

RESULT 8
US-10-264-237-2579
; Sequence 2579, Application US/10264237
; GENERAL INFORMATION:
; APPLICANT: Birse et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA131P1
; CURRENT APPLICATION NUMBER: US/10/264,237
; CURRENT FILING DATE: 2002-10-04
; PRIOR APPLICATION NUMBER: PCT/US01/16450
; PRIOR FILING DATE: 2001-05-18
; PRIOR APPLICATION NUMBER: US 60/205,515
; PRIOR FILING DATE: 2000-05-19
; NUMBER OF SEQ ID NOS: 2876
; SOFTWARE: PatentIn Ver. 3.1
; SEQ ID NO 2579
; LENGTH: 374
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-264-237-2579

Query Match 100.0%; Score 2000; DB 28; Length 374;
Best Local Similarity 100.0%; Pred. No. 3.4e-184;
Matches 374; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MRPGTALQAVLLAVLLVGLRAATGRLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAK 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MRPGTALQAVLLAVLLVGLRAATGRLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAK 60

Qy         61 EACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 EACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWT 120

Qy        121 DGSISQFRNWYVDEPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSD 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 DGSISQFRNWYVDEPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSD 180

Qy        181 EKPAVPSREAEGEETELTTPVLPEETQEEDAKKTFKESREAALNLAYILIPSIPLLLLLV 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 EKPAVPSREAEGEETELTTPVLPEETQEEDAKKTFKESREAALNLAYILIPSIPLLLLLV 240

Qy        241 VTTVVCWVWICRKRKREQPDPSTKKQHTIWPSPHQGNSPDLEVYNVIRKQSEADLAETRP 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 VTTVVCWVWICRKRKREQPDPSTKKQHTIWPSPHQGNSPDLEVYNVIRKQSEADLAETRP 300

Qy        301 DLKNISFRVCSGEATPDDMSCDYDNMAVNPSESGFVTLVSVESGFVTNDIYEFSPDQMGR 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 DLKNISFRVCSGEATPDDMSCDYDNMAVNPSESGFVTLVSVESGFVTNDIYEFSPDQMGR 360

Qy        361 SKESGWVENEIYGY 374
              ||||||||||||||
Db        361 SKESGWVENEIYGY 374
```

```
Db        361 SKESGWVENEIYGY 374


RESULT 9
US-10-648-593-148
; Sequence 148, Application US/10648593
; GENERAL INFORMATION:
;  APPLICANT: Bristol-Myers Squibb Company
;  TITLE OF INVENTION: IDENTIFICATION OF GENES FOR PREDICTING ACTIVITY OF COMPOUNDS THAT
;  TITLE OF INVENTION:  INTERACT WITH AND/OR MODULATE PROTEIN TYROSINE KINASES AND/OR
;  TITLE OF INVENTION:  PROTEIN TYROSINE KINASE PATHWAYS IN BREAST CELLS
;  FILE REFERENCE: D0273 NP
;  CURRENT APPLICATION NUMBER: US/10/648,593
;  CURRENT FILING DATE:  2003-08-26
;  PRIOR APPLICATION NUMBER: 60/406,385
;  PRIOR FILING DATE: 2002-08-27
;  NUMBER OF SEQ ID NOS: 557
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 148
;   LENGTH: 374
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-648-593-148

  Query Match           100.0%;  Score 2000;  DB 31;  Length 374;
  Best Local Similarity  100.0%;  Pred. No. 3.4e-184;
  Matches  374;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MRPGTALQAVLLAVLLVGLRAATGRLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAK 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MRPGTALQAVLLAVLLVGLRAATGRLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAK 60

Qy         61 EACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 EACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWT 120

Qy        121 DGSISQFRNWYVDEPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSD 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 DGSISQFRNWYVDEPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSD 180

Qy        181 EKPAVPSREAEGEETELTTPVLPEETQEEDAKKTFKESREAALNLAYILIPSIPLLLLLV 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 EKPAVPSREAEGEETELTTPVLPEETQEEDAKKTFKESREAALNLAYILIPSIPLLLLLV 240

Qy        241 VTTVVCWVWICRKRKREQPDPSTKKQHTIWPSPHQGNSPDLEVYNVIRKQSEADLAETRP 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 VTTVVCWVWICRKRKREQPDPSTKKQHTIWPSPHQGNSPDLEVYNVIRKQSEADLAETRP 300

Qy        301 DLKNISFRVCSGEATPDDMSCDYDNMAVNPSESGFVTLVSVESGFVTNDIYEFSPDQMGR 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 DLKNISFRVCSGEATPDDMSCDYDNMAVNPSESGFVTLVSVESGFVTNDIYEFSPDQMGR 360

Qy        361 SKESGWVENEIYGY 374
              ||||||||||||||
Db        361 SKESGWVENEIYGY 374


RESULT 10
US-60-172-354-15
; Sequence 15, Application US/60172354
; GENERAL INFORMATION:
;  APPLICANT: Yue, Henry
;  APPLICANT: Tang, Y. Tom
;  APPLICANT: Lal, Preeti
;  APPLICANT: Burford, Neil
;  APPLICANT: Azimzai, Yalda
;  APPLICANT: Patterson, Chandra
;  APPLICANT: Baughn, Mariah R.
;  APPLICANT: Lu, Dyung Aina M.
;  APPLICANT: Shah, Purvi
;  APPLICANT: Au-Young, Janice
;  TITLE OF INVENTION: EXTRACELLULAR MATRIX AND CELL ADHESION MOLECULES
;  FILE REFERENCE: PF-0760 P
;  CURRENT APPLICATION NUMBER: US/60/172,354
;  CURRENT FILING DATE: 1999-12-16
;  NUMBER OF SEQ ID NOS: 40
;  SOFTWARE: PERL Program
; SEQ ID NO 15
;   LENGTH: 374
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte Identifier No: 3143411CD1
US-60-172-354-15

  Query Match           100.0%;  Score 2000;  DB 33;  Length 374;
  Best Local Similarity  100.0%;  Pred. No. 3.4e-184;
  Matches  374;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MRPGTALQAVLLAVLLVGLRAATGRLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAK 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MRPGTALQAVLLAVLLVGLRAATGRLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAK 60

Qy         61 EACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 EACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWT 120

Qy        121 DGSISQFRNWYVDEPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSD 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 DGSISQFRNWYVDEPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSD 180

Qy        181 EKPAVPSREAEGEETELTTPVLPEETQEEDAKKTFKESREAALNLAYILIPSIPLLLLLV 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 EKPAVPSREAEGEETELTTPVLPEETQEEDAKKTFKESREAALNLAYILIPSIPLLLLLV 240

Qy        241 VTTVVCWVWICRKRKREQPDPSTKKQHTIWPSPHQGNSPDLEVYNVIRKQSEADLAETRP 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 VTTVVCWVWICRKRKREQPDPSTKKQHTIWPSPHQGNSPDLEVYNVIRKQSEADLAETRP 300

Qy        301 DLKNISFRVCSGEATPDDMSCDYDNMAVNPSESGFVTLVSVESGFVTNDIYEFSPDQMGR 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 DLKNISFRVCSGEATPDDMSCDYDNMAVNPSESGFVTLVSVESGFVTNDIYEFSPDQMGR 360

Qy        361 SKESGWVENEIYGY 374
              ||||||||||||||
Db        361 SKESGWVENEIYGY 374


RESULT 11
US-10-094-749-2142
; Sequence 2142, Application US/10094749
; GENERAL INFORMATION:
;  APPLICANT: ISOGAI, TAKAO
;  APPLICANT:  SUGIYAMA, TOMOYASU
;  APPLICANT:  OTSUKI, TETSUJI
;  APPLICANT:  WAKAMATSU, AI
;  APPLICANT:  SATO, HIROYUKI
;  APPLICANT:  ISHII, SHIZUKO
;  APPLICANT:  YAMAMOTO, JUN-ICHI
;  APPLICANT:  ISONO, YUUKO
;  APPLICANT:  HIO, YURI
;  APPLICANT:  OTSUKA, KAORU
;  APPLICANT:  NAGAI, KEIICHI
;  APPLICANT:  IRIE, RYOTARO
;  APPLICANT:  TAMECHIKA, ICHIRO
;  APPLICANT:  SEKI, NAOHIKO
;  APPLICANT:  YOSHIKAWA, TSUTOMU
;  APPLICANT:  OTSUKA, MOTOYUKI
;  APPLICANT:  NAGAHARI, KENJI
;  APPLICANT:  MASUHO, YASUHIKO
;  TITLE OF INVENTION: NOVEL FULL-LENGTH cDNA
;  FILE REFERENCE: 084335/0160
;  CURRENT APPLICATION NUMBER: US/10/094,749
;  CURRENT FILING DATE:  2002-03-12
;  PRIOR APPLICATION NUMBER: 60/350,435
```

; PRIOR FILING DATE: 2002-01-24
; PRIOR APPLICATION NUMBER: JP 2001-328381
; PRIOR FILING DATE: 2001-09-14
; NUMBER OF SEQ ID NOS: 3381
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2142
; LENGTH: 374
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-094-749-2142

Query Match 99.8%; Score 1996; DB 26; Length 374;
Best Local Similarity 99.7%; Pred. No. 8.4e-184;
Matches 373; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 MRPGTALQAVLLAVLLVGLRAATGRLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAK 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MRPGTALQAVLLAVLLVGLRAATGRLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAK 60

Qy         61 EACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 EACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWT 120

Qy        121 DGSISQFRNWYVDEPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSD 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 DGSISQFRNWYVDEPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSD 180

Qy        181 EKPAVPSREAEGEETELTTPVLPEETQEEDAKKTFKESREAALNLAYILIPSIPLLLLLV 240
              |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db        181 EKPAVPSREAEGEETELTTPVLPEETQEEDTKKTFKESREAALNLAYILIPSIPLLLLLV 240

Qy        241 VTTVVCWVWICRKRKREQPDPSTKKQHTIWPSPHQGNSPDLEVYNVIRKQSEADLAETRP 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 VTTVVCWVWICRKRKREQPDPSTKKQHTIWPSPHQGNSPDLEVYNVIRKQSEADLAETRP 300

Qy        301 DLKNISFRVCSGEATPDDMSCDYDNMAVNPSESGFVTLVSVESGFVTNDIYEFSPDQMGR 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 DLKNISFRVCSGEATPDDMSCDYDNMAVNPSESGFVTLVSVESGFVTNDIYEFSPDQMGR 360

Qy        361 SKESGWVENEIYGY 374
              ||||||||||||||
Db        361 SKESGWVENEIYGY 374
```

RESULT 12
PCT-US99-17130-163
; Sequence 163, Application PC/TUS9917130
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
; TITLE OF INVENTION: 98 Human Secreted Proteins
; FILE REFERENCE: PZ031.PCT
; CURRENT APPLICATION NUMBER: PCT/US99/17130
; CURRENT FILING DATE: 1999-07-29
; EARLIER APPLICATION NUMBER: 60/094,657
; EARLIER FILING DATE: 1998-07-30
; EARLIER APPLICATION NUMBER: 60/095,486
; EARLIER FILING DATE: 1998-08-05
; EARLIER APPLICATION NUMBER: 60/095,455
; EARLIER FILING DATE: 1998-08-06
; EARLIER APPLICATION NUMBER: 60/095,454
; EARLIER FILING DATE: 1998-08-06
; EARLIER APPLICATION NUMBER: 60/096,319
; EARLIER FILING DATE: 1998-08-12
; NUMBER OF SEQ ID NOS: 364
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 163
; LENGTH: 374
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (84)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (112)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
PCT-US99-17130-163

Query Match 99.5%; Score 1990; DB 1; Length 374;
Best Local Similarity 99.5%; Pred. No. 3.2e-183;
Matches 372; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy          1 MRPGTALQAVLLAVLLVGLRAATGRLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAK 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MRPGTALQAVLLAVLLVGLRAATGRLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAK 60

Qy         61 EACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWT 120
              |||||||||||||||||||||||| |||||||||||||||||||||||||| ||||||||
Db         61 EACRRDGGQLVSIESEDEQKLIEXFIENLLPSDGDFWIGLRRREEKQSNSTXCQDLYAWT 120

Qy        121 DGSISQFRNWYVDEPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSD 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 DGSISQFRNWYVDEPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSD 180

Qy        181 EKPAVPSREAEGEETELTTPVLPEETQEEDAKKTFKESREAALNLAYILIPSIPLLLLLV 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 EKPAVPSREAEGEETELTTPVLPEETQEEDAKKTFKESREAALNLAYILIPSIPLLLLLV 240

Qy        241 VTTVVCWVWICRKRKREQPDPSTKKQHTIWPSPHQGNSPDLEVYNVIRKQSEADLAETRP 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 VTTVVCWVWICRKRKREQPDPSTKKQHTIWPSPHQGNSPDLEVYNVIRKQSEADLAETRP 300

Qy        301 DLKNISFRVCSGEATPDDMSCDYDNMAVNPSESGFVTLVSVESGFVTNDIYEFSPDQMGR 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 DLKNISFRVCSGEATPDDMSCDYDNMAVNPSESGFVTLVSVESGFVTNDIYEFSPDQMGR 360

Qy        361 SKESGWVENEIYGY 374
              ||||||||||||||
Db        361 SKESGWVENEIYGY 374
```

RESULT 13
US-10-351-334-166
; Sequence 166, Application US/10351334
; GENERAL INFORMATION:
; APPLICANT: Komatsoulis et al
; TITLE OF INVENTION: 98 Human Secreted Proteins
; FILE REFERENCE: PZ031P2
; CURRENT APPLICATION NUMBER: US/10/351,334
; CURRENT FILING DATE: 2003-01-27
; PRIOR APPLICATION NUMBER: 60/350,898
; PRIOR FILING DATE: 2002-01-25
; PRIOR APPLICATION NUMBER: 09/489,847
; PRIOR FILING DATE: 2000-01-24
; PRIOR APPLICATION NUMBER: PCT/US99/17130
; PRIOR FILING DATE: 1999-07-29
; PRIOR APPLICATION NUMBER: 60/094,657
; PRIOR FILING DATE: 1998-07-30
; PRIOR APPLICATION NUMBER: 60/095,486
; PRIOR FILING DATE: 1998-08-05
; PRIOR APPLICATION NUMBER: 60/096,319
; PRIOR FILING DATE: 1998-08-12
; PRIOR APPLICATION NUMBER: 60/095,454
; PRIOR FILING DATE: 1998-08-06
; PRIOR APPLICATION NUMBER: 60/095,455
; PRIOR FILING DATE: 1998-08-06
; NUMBER OF SEQ ID NOS: 376
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 166
; LENGTH: 374
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (84)

```
;   OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
;   FEATURE:
;   NAME/KEY: SITE
;   LOCATION: (112)
;   OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-351-334-166

  Query Match           99.5%;  Score 1990;  DB 29;  Length 374;
  Best Local Similarity  99.5%;  Pred. No. 3.2e-183;
  Matches  372;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 MRPGTALQAVLLAVLLVGLRAATGRLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAK 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MRPGTALQAVLLAVLLVGLRAATGRLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAK 60

Qy         61 EACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWT 120
              |||||||||||||||||||| |||||||||||||||||||||||| ||||||||
Db         61 EACRRDGGQLVSIESEDEQKLIEXFIENLLPSDGDFWIGLRRREEKQSNSTXCQDLYAWT 120

Qy        121 DGSISQFRNWYVDEPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSD 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 DGSISQFRNWYVDEPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSD 180

Qy        181 EKPAVPSREAEGEETELTTPVLPEETQEEDAKKTFKESREAALNLAYILIPSIPLLLLLV 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 EKPAVPSREAEGEETELTTPVLPEETQEEDAKKTFKESREAALNLAYILIPSIPLLLLLV 240

Qy        241 VTTVVCWVWICRKRKREQPDPSTKKQHTIWPSPHQGNSPDLEVYNVIRKQSEADLAETRP 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 VTTVVCWVWICRKRKREQPDPSTKKQHTIWPSPHQGNSPDLEVYNVIRKQSEADLAETRP 300

Qy        301 DLKNISFRVCSGEATPDDMSCDYDNMAVNPSESGFVTLVSVESGFVTNDIYEFSPDQMGR 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 DLKNISFRVCSGEATPDDMSCDYDNMAVNPSESGFVTLVSVESGFVTNDIYEFSPDQMGR 360

Qy        361 SKESGWVENEIYGY 374
              ||||||||||||||
Db        361 SKESGWVENEIYGY 374


RESULT 14
US-09-380-139A-137
; Sequence 137, Application US/09380139A
; GENERAL INFORMATION:
;  APPLICANT: GENENTECH, INC. et al.
;  APPLICANT:  Chen, Jian
;  APPLICANT:  Goddard, Audrey
;  APPLICANT:  Gurney, Austin L.
;  APPLICANT:  Pennica, Diane
;  APPLICANT:  Wood, William I.
;  APPLICANT:  Yuan, Jean
;  TITLE OF INVENTION: NOVEL POLYPEPTIDES AND NUCLEIC ACIDS ENCODING THE SAME
;  FILE REFERENCE: 10466-04
;  CURRENT APPLICATION NUMBER: US/09/380,139A
;  CURRENT FILING DATE:  1999-08-25
;  PRIOR APPLICATION NUMBER: PCT/US 98/19330
;  PRIOR FILING DATE: 1998-09-16
;  NUMBER OF SEQ ID NOS: 379
; SEQ ID NO 137
;   LENGTH: 382
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-380-139A-137

  Query Match           99.3%;  Score 1986;  DB 17;  Length 382;
  Best Local Similarity  97.9%;  Pred. No. 8.1e-183;
  Matches  374;  Conservative    0;  Mismatches    0;  Indels    8;  Gaps    1;

Qy          1 MRPGTALQAVLLAVLLVGLRAATGRLLS--------GQPVCRGGTQRPCYKVIYFHDTSR 52
              ||||||||||||||||||||||||||||        ||||||||||||||||||||||||
Db          1 MRPGTALQAVLLAVLLVGLRAATGRLLSASDLDLRGGQPVCRGGTQRPCYKVIYFHDTSR 60

Qy         53 RLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTA 112
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 RLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTA 120

Qy        113 CQDLYAWTDGSISQFRNWYVDEPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKN 172
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CQDLYAWTDGSISQFRNWYVDEPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKN 180

Qy        173 NFICKYSDEKPAVPSREAEGEETELTTPVLPEETQEEDAKKTFKESREAALNLAYILIPS 232
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 NFICKYSDEKPAVPSREAEGEETELTTPVLPEETQEEDAKKTFKESREAALNLAYILIPS 240

Qy        233 IPLLLLLVVTTVVCWVWICRKRKREQPDPSTKKQHTIWPSPHQGNSPDLEVYNVIRKQSE 292
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 IPLLLLLVVTTVVCWVWICRKRKREQPDPSTKKQHTIWPSPHQGNSPDLEVYNVIRKQSE 300

Qy        293 ADLAETRPDLKNISFRVCSGEATPDDMSCDYDNMAVNPSESGFVTLVSVESGFVTNDIYE 352
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ADLAETRPDLKNISFRVCSGEATPDDMSCDYDNMAVNPSESGFVTLVSVESGFVTNDIYE 360

Qy        353 FSPDQMGRSKESGWVENEIYGY 374
              ||||||||||||||||||||||
Db        361 FSPDQMGRSKESGWVENEIYGY 382


RESULT 15
US-09-423-844-137
; Sequence 137, Application US/09423844
; GENERAL INFORMATION:
;  APPLICANT: Genentech, Inc.
;  APPLICANT:  Chen, Jian
;  APPLICANT:  Goddard, Audrey
;  APPLICANT:  Gurney, Austin L.
;  APPLICANT:  Hillan, Kenneth
;  APPLICANT:  Pennica, Diane
;  APPLICANT:  Wood, William I.
;  APPLICANT:  Yuan, Jean
;  TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
;  TITLE OF INVENTION:  Acids Encoding the Same
;  FILE REFERENCE: P1618P1E
;  CURRENT APPLICATION NUMBER: US/09/423,844
;  CURRENT FILING DATE:  1999-11-12
;  PRIOR APPLICATION NUMBER: PCT/US99/21090
;  PRIOR FILING DATE: 1999-09-15
;  PRIOR APPLICATION NUMBER: PCT/US98/19330
;  PRIOR FILING DATE: 1998-09-16
;  NUMBER OF SEQ ID NOS: 379
; SEQ ID NO 137
;   LENGTH: 382
;   TYPE: PRT
;   ORGANISM: Homo Sapien
US-09-423-844-137

  Query Match           99.3%;  Score 1986;  DB 18;  Length 382;
  Best Local Similarity  97.9%;  Pred. No. 8.1e-183;
  Matches  374;  Conservative    0;  Mismatches    0;  Indels    8;  Gaps    1;

Qy          1 MRPGTALQAVLLAVLLVGLRAATGRLLS--------GQPVCRGGTQRPCYKVIYFHDTSR 52
              ||||||||||||||||||||||||||||        ||||||||||||||||||||||||
Db          1 MRPGTALQAVLLAVLLVGLRAATGRLLSASDLDLRGGQPVCRGGTQRPCYKVIYFHDTSR 60

Qy         53 RLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTA 112
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 RLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTA 120

Qy        113 CQDLYAWTDGSISQFRNWYVDEPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKN 172
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CQDLYAWTDGSISQFRNWYVDEPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKN 180

Qy        173 NFICKYSDEKPAVPSREAEGEETELTTPVLPEETQEEDAKKTFKESREAALNLAYILIPS 232
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 NFICKYSDEKPAVPSREAEGEETELTTPVLPEETQEEDAKKTFKESREAALNLAYILIPS 240

Qy        233 IPLLLLLVVTTVVCWVWICRKRKREQPDPSTKKQHTIWPSPHQGNSPDLEVYNVIRKQSE 292
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 IPLLLLLVVTTVVCWVWICRKRKREQPDPSTKKQHTIWPSPHQGNSPDLEVYNVIRKQSE 300

Qy        293 ADLAETRPDLKNISFRVCSGEATPDDMSCDYDNMAVNPSESGFVTLVSVESGFVTNDIYE 352
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ADLAETRPDLKNISFRVCSGEATPDDMSCDYDNMAVNPSESGFVTLVSVESGFVTNDIYE 360

Qy        353 FSPDQMGRSKESGWVENEIYGY 374
                ||||||||||||||||||||||
Db        361 FSPDQMGRSKESGWVENEIYGY 382
```

Search completed: September 9, 2004, 22:30:43
Job time : 419 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: September 9, 2004, 22:38:59 ; Search time 128 Seconds
(without alignments)
516.110 Million cell updates/sec

Title: US-09-887-855-2_COPY_22_227
Perfect score: 1115
Sequence: 1 ATGRLLSGQPVCRGGTQRPC.....EEDAKKTFKESREAALNLAY 206

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1335176 seqs, 320689617 residues

Total number of hits satisfying chosen parameters: 1335176

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 50 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1115	100.0	206	9	US-09-887-855-5	Sequence 5, Appli
2	1115	100.0	374	9	US-09-887-855-2	Sequence 2, Appli
3	1115	100.0	374	14	US-10-149-819-15	Sequence 15, Appl
4	1115	100.0	374	15	US-10-094-749-2090	Sequence 2090, Ap
5	1115	100.0	374	15	US-10-264-237-2579	Sequence 2579, Ap
6	1115	100.0	374	16	US-10-648-593-148	Sequence 148, App
7	1115	100.0	387	12	US-10-296-115-1311	Sequence 1311, Ap
8	1111	99.6	374	15	US-10-094-749-2142	Sequence 2142, Ap
9	1105	99.1	374	12	US-10-351-334-166	Sequence 166, App
10	1101	98.7	382	9	US-09-909-320-137	Sequence 137, App
11	1101	98.7	382	9	US-09-909-088B-137	Sequence 137, App
12	1101	98.7	382	9	US-09-905-291A-137	Sequence 137, App
13	1101	98.7	382	9	US-09-902-853-137	Sequence 137, App
14	1101	98.7	382	9	US-09-907-824-137	Sequence 137, App
15	1101	98.7	382	9	US-09-907-841-137	Sequence 137, App
16	1101	98.7	382	10	US-09-904-011-137	Sequence 137, App
17	1101	98.7	382	10	US-09-906-742-137	Sequence 137, App
18	1101	98.7	382	10	US-09-906-838-137	Sequence 137, App
19	1101	98.7	382	10	US-09-907-613-137	Sequence 137, App
20	1101	98.7	382	10	US-09-907-942-137	Sequence 137, App
21	1101	98.7	382	10	US-09-904-859-137	Sequence 137, App
22	1101	98.7	382	10	US-09-909-204-137	Sequence 137, App
23	1101	98.7	382	10	US-09-904-820-137	Sequence 137, App
24	1101	98.7	382	10	US-09-904-786-137	Sequence 137, App
25	1101	98.7	382	10	US-09-906-646-137	Sequence 137, App
26	1101	98.7	382	10	US-09-906-700-137	Sequence 137, App
27	1101	98.7	382	10	US-09-903-786-137	Sequence 137, App
28	1101	98.7	382	10	US-09-902-903-137	Sequence 137, App
29	1101	98.7	382	10	US-09-903-749A-137	Sequence 137, App
30	1101	98.7	382	10	US-09-904-119-137	Sequence 137, App
31	1101	98.7	382	10	US-09-904-956-137	Sequence 137, App
32	1101	98.7	382	10	US-09-902-736-137	Sequence 137, App
33	1101	98.7	382	10	US-09-907-794-137	Sequence 137, App
34	1101	98.7	382	10	US-09-903-943-137	Sequence 137, App
35	1101	98.7	382	10	US-09-904-462-137	Sequence 137, App
36	1101	98.7	382	10	US-09-907-925-137	Sequence 137, App
37	1101	98.7	382	10	US-09-902-692-137	Sequence 137, App
38	1101	98.7	382	10	US-09-903-520-137	Sequence 137, App
39	1101	98.7	382	10	US-09-905-056-137	Sequence 137, App
40	1101	98.7	382	10	US-09-909-064-137	Sequence 137, App
41	1101	98.7	382	10	US-09-904-553-137	Sequence 137, App
42	1101	98.7	382	10	US-09-905-381-137	Sequence 137, App
43	1101	98.7	382	10	US-09-905-088-137	Sequence 137, App
44	1101	98.7	382	10	US-09-907-575-137	Sequence 137, App
45	1101	98.7	382	10	US-09-905-075-137	Sequence 137, App
46	1101	98.7	382	10	US-09-902-759-137	Sequence 137, App
47	1101	98.7	382	10	US-09-902-634-137	Sequence 137, App
48	1101	98.7	382	10	US-09-902-713-137	Sequence 137, App
49	1101	98.7	382	10	US-09-907-979-137	Sequence 137, App
50	1101	98.7	382	10	US-09-902-615-137	Sequence 137, App

ALIGNMENTS

RESULT 1
US-09-887-855-5
; Sequence 5, Application US/09887855
; Patent No. US20020058310A1
; GENERAL INFORMATION:
; APPLICANT: Immunex Corporation
; APPLICANT: Anderson, Dirk M
; TITLE OF INVENTION: LECTIN SS3939 DNA AND POLYPEPTIDES
; FILE REFERENCE: 2883-US
; CURRENT APPLICATION NUMBER: US/09/887,855
; CURRENT FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 206
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-887-855-5

Query Match 100.0%; Score 1115; DB 9; Length 206;
Best Local Similarity 100.0%; Pred. No. 1.5e-104;
Matches 206; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 ATGRLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKL 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGRLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKL 60

Qy         61 IEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEPSCGSEV 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 IEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEPSCGSEV 120

Qy        121 CVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGEETELTTPV 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        121 CVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGEETELTTPV 180

Qy        181 LPEETQEEDAKKTFKESREAALNLAY 206
              |||||||||||||||||||||||||||
Db        181 LPEETQEEDAKKTFKESREAALNLAY 206


RESULT 2
US-09-887-855-2
; Sequence 2, Application US/09887855
; Patent No. US20020058310A1
; GENERAL INFORMATION:
;  APPLICANT: Immunex Corporation
;  APPLICANT:  Anderson, Dirk M
;  TITLE OF INVENTION: LECTIN SS3939 DNA AND POLYPEPTIDES
;  FILE REFERENCE: 2883-US
;  CURRENT APPLICATION NUMBER: US/09/887,855
;  CURRENT FILING DATE:  2001-06-22
;  NUMBER OF SEQ ID NOS: 9
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
;   LENGTH: 374
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-887-855-2

  Query Match             100.0%;  Score 1115;  DB 9;  Length 374;
  Best Local Similarity   100.0%;  Pred. No. 3.3e-104;
  Matches  206;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ATGRLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKL 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ATGRLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKL 81

Qy         61 IEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEPSCGSEV 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 IEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEPSCGSEV 141

Qy        121 CVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGEETELTTPV 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 CVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGEETELTTPV 201

Qy        181 LPEETQEEDAKKTFKESREAALNLAY 206
              |||||||||||||||||||||||||||
Db        202 LPEETQEEDAKKTFKESREAALNLAY 227


RESULT 3
US-10-149-819-15
; Sequence 15, Application US/10149819
; Publication No. US20030044913A1
; GENERAL INFORMATION:
;  APPLICANT: INCYTE GENOMICS, INC.
;  APPLICANT:  YUE, Henry
;  APPLICANT:  AZIMZAI, Yalda
;  APPLICANT:  TANG, Y. Tom
;  APPLICANT:  PATTERSON, Chandra
;  APPLICANT:  BAUGHN, Mariah R.
;  APPLICANT:  LU, Dyung Aina M.
;  APPLICANT:  SHAH, Purvi
;  APPLICANT:  LAL, Preeti
;  APPLICANT:  AU-YOUNG, Janice
;  APPLICANT:  BURFORD, Neil
;  TITLE OF INVENTION: EXTRACELLULAR MATRIX AND CELL ADHESION MOLECULES
;  FILE REFERENCE: PF-0760 PCT
;  CURRENT APPLICATION NUMBER: US/10/149,819
;  CURRENT FILING DATE:  2002-06-10
;  PRIOR APPLICATION NUMBER: 60/172,852; 60/172,354
;  PRIOR FILING DATE: 1999-12-10; 1999-12-16
;  NUMBER OF SEQ ID NOS: 42
;  SOFTWARE: PERL Program
; SEQ ID NO 15
;   LENGTH: 374
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No. US20030044913A1 3143411CD1
US-10-149-819-15

  Query Match             100.0%;  Score 1115;  DB 14;  Length 374;
  Best Local Similarity   100.0%;  Pred. No. 3.3e-104;
  Matches  206;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ATGRLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKL 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ATGRLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKL 81

Qy         61 IEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEPSCGSEV 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 IEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEPSCGSEV 141

Qy        121 CVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGEETELTTPV 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 CVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGEETELTTPV 201

Qy        181 LPEETQEEDAKKTFKESREAALNLAY 206
              |||||||||||||||||||||||||||
Db        202 LPEETQEEDAKKTFKESREAALNLAY 227


RESULT 4
US-10-094-749-2090
; Sequence 2090, Application US/10094749
; Publication No. US20030219741A1
; GENERAL INFORMATION:
;  APPLICANT: ISOGAI, TAKAO
;  APPLICANT:  SUGIYAMA, TOMOYASU
;  APPLICANT:  OTSUKI, TETSUJI
;  APPLICANT:  WAKAMATSU, AI
;  APPLICANT:  SATO, HIROYUKI
;  APPLICANT:  ISHII, SHIZUKO
;  APPLICANT:  YAMAMOTO, JUN-ICHI
;  APPLICANT:  ISONO, YUUKO
;  APPLICANT:  HIO, YURI
;  APPLICANT:  OTSUKA, KAORU
;  APPLICANT:  NAGAI, KEIICHI
;  APPLICANT:  IRIE, RYOTARO
;  APPLICANT:  TAMECHIKA, ICHIRO
;  APPLICANT:  SEKI, NAOHIKO
;  APPLICANT:  YOSHIKAWA, TSUTOMU
;  APPLICANT:  OTSUKA, MOTOYUKI
;  APPLICANT:  NAGAHARI, KENJI
;  APPLICANT:  MASUHO, YASUHIKO
;  TITLE OF INVENTION: NOVEL FULL-LENGTH cDNA
;  FILE REFERENCE: 084335/0160
;  CURRENT APPLICATION NUMBER: US/10/094,749
;  CURRENT FILING DATE:  2002-03-12
;  PRIOR APPLICATION NUMBER: 60/350,435
;  PRIOR FILING DATE: 2002-01-24
;  PRIOR APPLICATION NUMBER: JP 2001-328381
;  PRIOR FILING DATE: 2001-09-14
;  NUMBER OF SEQ ID NOS: 3381
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2090
;   LENGTH: 374
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-094-749-2090

  Query Match             100.0%;  Score 1115;  DB 15;  Length 374;
  Best Local Similarity   100.0%;  Pred. No. 3.3e-104;
  Matches  206;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ATGRLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKL 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db         22 ATGRLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKL 81

Qy         61 IEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEPSCGSEV 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 IEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEPSCGSEV 141

Qy        121 CVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGEETELTTPV 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 CVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGEETELTTPV 201

Qy        181 LPEETQEEDAKKTFKESREAALNLAY 206
              |||||||||||||||||||||||||
Db        202 LPEETQEEDAKKTFKESREAALNLAY 227
```

RESULT 5
US-10-264-237-2579
; Sequence 2579, Application US/10264237
; Publication No. US20040009491A1
; GENERAL INFORMATION:
; APPLICANT: Birse et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA131P1
; CURRENT APPLICATION NUMBER: US/10/264,237
; CURRENT FILING DATE: 2002-10-04
; PRIOR APPLICATION NUMBER: PCT/US01/16450
; PRIOR FILING DATE: 2001-05-18
; PRIOR APPLICATION NUMBER: US 60/205,515
; PRIOR FILING DATE: 2000-05-19
; NUMBER OF SEQ ID NOS: 2876
; SOFTWARE: PatentIn Ver. 3.1
; SEQ ID NO 2579
; LENGTH: 374
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-264-237-2579

Query Match 100.0%; Score 1115; DB 15; Length 374;
Best Local Similarity 100.0%; Pred. No. 3.3e-104;
Matches 206; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 ATGRLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKL 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ATGRLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKL 81

Qy         61 IEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEPSCGSEV 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 IEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEPSCGSEV 141

Qy        121 CVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGEETELTTPV 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 CVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGEETELTTPV 201

Qy        181 LPEETQEEDAKKTFKESREAALNLAY 206
              |||||||||||||||||||||||||
Db        202 LPEETQEEDAKKTFKESREAALNLAY 227
```

RESULT 6
US-10-648-593-148
; Sequence 148, Application US/10648593
; Publication No. US20040106132A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: IDENTIFICATION OF GENES FOR PREDICTING ACTIVITY OF COMPOUNDS THAT
; TITLE OF INVENTION: INTERACT WITH AND/OR MODULATE PROTEIN TYROSINE KINASES AND/OR
; TITLE OF INVENTION: PROTEIN TYROSINE KINASE PATHWAYS IN BREAST CELLS
; FILE REFERENCE: D0273 NP
; CURRENT APPLICATION NUMBER: US/10/648,593
; CURRENT FILING DATE: 2003-08-26
; PRIOR APPLICATION NUMBER: 60/406,385
; PRIOR FILING DATE: 2002-08-27
; NUMBER OF SEQ ID NOS: 557
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 148
; LENGTH: 374
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-648-593-148

Query Match 100.0%; Score 1115; DB 16; Length 374;
Best Local Similarity 100.0%; Pred. No. 3.3e-104;
Matches 206; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 ATGRLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKL 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ATGRLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKL 81

Qy         61 IEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEPSCGSEV 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 IEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEPSCGSEV 141

Qy        121 CVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGEETELTTPV 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 CVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGEETELTTPV 201

Qy        181 LPEETQEEDAKKTFKESREAALNLAY 206
              |||||||||||||||||||||||||
Db        202 LPEETQEEDAKKTFKESREAALNLAY 227
```

RESULT 7
US-10-296-115-1311
; Sequence 1311, Application US/10296115
; Publication No. US20040053248A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq Inc
; TITLE OF INVENTION: No. US20040053248A1el Nucleic Acids and Polypeptides
; FILE REFERENCE: 784PCT
; CURRENT APPLICATION NUMBER: US/10/296,115
; CURRENT FILING DATE: 2002-11-18
; PRIOR APPLICATION NUMBER: US09/488,725
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: US09/552,317
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 1478
; SEQ ID NO 1311
; LENGTH: 387
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-296-115-1311

Query Match 100.0%; Score 1115; DB 12; Length 387;
Best Local Similarity 100.0%; Pred. No. 3.4e-104;
Matches 206; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 ATGRLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKL 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         35 ATGRLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKL 94

Qy         61 IEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEPSCGSEV 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         95 IEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEPSCGSEV 154

Qy        121 CVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGEETELTTPV 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        155 CVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGEETELTTPV 214

Qy        181 LPEETQEEDAKKTFKESREAALNLAY 206
              |||||||||||||||||||||||||
Db        215 LPEETQEEDAKKTFKESREAALNLAY 240
```

RESULT 8
US-10-094-749-2142
; Sequence 2142, Application US/10094749

; Publication No. US20030219741A1
; GENERAL INFORMATION:
; APPLICANT: ISOGAI, TAKAO
; APPLICANT: SUGIYAMA, TOMOYASU
; APPLICANT: OTSUKI, TETSUJI
; APPLICANT: WAKAMATSU, AI
; APPLICANT: SATO, HIROYUKI
; APPLICANT: ISHII, SHIZUKO
; APPLICANT: YAMAMOTO, JUN-ICHI
; APPLICANT: ISONO, YUUKO
; APPLICANT: HIO, YURI
; APPLICANT: OTSUKA, KAORU
; APPLICANT: NAGAI, KEIICHI
; APPLICANT: IRIE, RYOTARO
; APPLICANT: TAMECHIKA, ICHIRO
; APPLICANT: SEKI, NAOHIKO
; APPLICANT: YOSHIKAWA, TSUTOMU
; APPLICANT: OTSUKA, MOTOYUKI
; APPLICANT: NAGAHARI, KENJI
; APPLICANT: MASUHO, YASUHIKO
; TITLE OF INVENTION: NOVEL FULL-LENGTH cDNA
; FILE REFERENCE: 084335/0160
; CURRENT APPLICATION NUMBER: US/10/094,749
; CURRENT FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 60/350,435
; PRIOR FILING DATE: 2002-01-24
; PRIOR APPLICATION NUMBER: JP 2001-328381
; PRIOR FILING DATE: 2001-09-14
; NUMBER OF SEQ ID NOS: 3381
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2142
; LENGTH: 374
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-094-749-2142

Query Match 99.6%; Score 1111; DB 15; Length 374;
Best Local Similarity 99.5%; Pred. No. 8.3e-104;
Matches 205; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 ATGRLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKL 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ATGRLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKL 81

Qy         61 IEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEPSCGSEV 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 IEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEPSCGSEV 141

Qy        121 CVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGEETELTTPV 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 CVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGEETELTTPV 201

Qy        181 LPEETQEEDAKKTFKESREAALNLAY 206
              |||||||| ||||||||||||||||
Db        202 LPEETQEEDTKKTFKESREAALNLAY 227
```

RESULT 9
US-10-351-334-166
; Sequence 166, Application US/10351334
; Publication No. US20040034196A1
; GENERAL INFORMATION:
; APPLICANT: Komatsoulis et al
; TITLE OF INVENTION: 98 Human Secreted Proteins
; FILE REFERENCE: PZ031P2
; CURRENT APPLICATION NUMBER: US/10/351,334
; CURRENT FILING DATE: 2003-01-27
; PRIOR APPLICATION NUMBER: 60/350,898
; PRIOR FILING DATE: 2002-01-25
; PRIOR APPLICATION NUMBER: 09/489,847
; PRIOR FILING DATE: 2000-01-24
; PRIOR APPLICATION NUMBER: PCT/US99/17130
; PRIOR FILING DATE: 1999-07-29
; PRIOR APPLICATION NUMBER: 60/094,657
; PRIOR FILING DATE: 1998-07-30
; PRIOR APPLICATION NUMBER: 60/095,486
; PRIOR FILING DATE: 1998-08-05
; PRIOR APPLICATION NUMBER: 60/096,319
; PRIOR FILING DATE: 1998-08-12
; PRIOR APPLICATION NUMBER: 60/095,454
; PRIOR FILING DATE: 1998-08-06
; PRIOR APPLICATION NUMBER: 60/095,455
; PRIOR FILING DATE: 1998-08-06
; NUMBER OF SEQ ID NOS: 376
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 166
; LENGTH: 374
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (84)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (112)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-351-334-166

Query Match 99.1%; Score 1105; DB 12; Length 374;
Best Local Similarity 99.0%; Pred. No. 3.4e-103;
Matches 204; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy          1 ATGRLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKL 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ATGRLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKL 81

Qy         61 IEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEPSCGSEV 120
              || |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db         82 IEXFIENLLPSDGDFWIGLRRREEKQSNSTXCQDLYAWTDGSISQFRNWYVDEPSCGSEV 141

Qy        121 CVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGEETELTTPV 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 CVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGEETELTTPV 201

Qy        181 LPEETQEEDAKKTFKESREAALNLAY 206
              ||||||||||||||||||||||||||
Db        202 LPEETQEEDAKKTFKESREAALNLAY 227
```

RESULT 10
US-09-909-320-137
; Sequence 137, Application US/09909320
; Patent No. US20020132240A1
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann

```
;  APPLICANT:  Stewart, Timothy A.
;  APPLICANT:  Tumas, Daniel
;  APPLICANT:  Williams, P. Mickey
;  APPLICANT:  Wood, William, I.
;  TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
;  TITLE OF INVENTION:  Acids Encoding the Same
;  FILE REFERENCE: 10466-14
;  CURRENT APPLICATION NUMBER: US/09/909,320
;  CURRENT FILING DATE:  2002-01-04
;  PRIOR APPLICATION NUMBER: PCT/US00/04414
;  PRIOR FILING DATE: 2000-02-22
;  PRIOR APPLICATION NUMBER: US 60/143,048
;  PRIOR FILING DATE: 1999-07-07
;  PRIOR APPLICATION NUMBER: US 60/145,698
;  PRIOR FILING DATE: 1999-07-26
;  PRIOR APPLICATION NUMBER: US 60/146,222
;  PRIOR FILING DATE: 1999-07-28
;  PRIOR APPLICATION NUMBER: PCT/US99/20594
;  PRIOR FILING DATE: 1999-09-08
;  PRIOR APPLICATION NUMBER: PCT/US99/20944
;  PRIOR FILING DATE: 1999-09-13
;  PRIOR APPLICATION NUMBER: PCT/US99/21090
;  PRIOR FILING DATE: 1999-09-15
;  PRIOR APPLICATION NUMBER: PCT/US99/21547
;  PRIOR FILING DATE: 1999-09-15
;  PRIOR APPLICATION NUMBER: PCT/US99/23089
;  PRIOR FILING DATE: 1999-10-05
;  PRIOR APPLICATION NUMBER: PCT/US99/28214
;  PRIOR FILING DATE: 1999-11-29
;  PRIOR APPLICATION NUMBER: PCT/US99/28313
;  PRIOR FILING DATE: 1999-11-30
;  PRIOR APPLICATION NUMBER: PCT/US99/28564
;  PRIOR FILING DATE: 1999-12-02
;  PRIOR APPLICATION NUMBER: PCT/US99/28565
;  PRIOR FILING DATE: 1999-12-02
;  PRIOR APPLICATION NUMBER: PCT/US99/30095
;  PRIOR FILING DATE: 1999-12-16
;  PRIOR APPLICATION NUMBER: PCT/US99/30911
;  PRIOR FILING DATE: 1999-12-20
;  PRIOR APPLICATION NUMBER: PCT/US99/30999
;  PRIOR FILING DATE: 1999-12-20
;  PRIOR APPLICATION NUMBER: PCT/US00/00219
;  PRIOR FILING DATE: 2000-01-05
;  NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 137
;   LENGTH: 382
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-909-320-137

  Query Match             98.7%;  Score 1101;  DB 9;  Length 382;
  Best Local Similarity   96.3%;  Pred. No. 8.8e-103;
  Matches  206;  Conservative    0;  Mismatches    0;  Indels    8;  Gaps    1;

Qy          1 ATGRLLS--------GQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 52
              |||||||        ||||||||||||||||||||||||||||||||||||||||||||
Db         22 ATGRLLSASDLDLRGGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 81

Qy         53 ESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVD 112
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 ESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVD 141

Qy        113 EPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGE 172
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 EPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGE 201

Qy        173 ETELTTPVLPEETQEEDAKKTFKESREAALNLAY 206
              |||||||||||||||||||||||||||||||||
Db        202 ETELTTPVLPEETQEEDAKKTFKESREAALNLAY 235


RESULT 11
US-09-909-088B-137
; Sequence 137, Application US/09909088B
; Patent No. US20020146709A1
; GENERAL INFORMATION:
;  APPLICANT: Genentech, Inc.
;  APPLICANT:  Ashkenazi, Avi
;  APPLICANT:  Botstein, David
;  APPLICANT:  Desnoyers, Luc
;  APPLICANT:  Eaton, Dan L.
;  APPLICANT:  Ferrara, Napoleone
;  APPLICANT:  Filvaroff, Ellen
;  APPLICANT:  Fong, Sherman
;  APPLICANT:  Gao, Wei-Qiang
;  APPLICANT:  Gerber, Hanspeter
;  APPLICANT:  Gerritsen, Mary E.
;  APPLICANT:  Goddard, A.
;  APPLICANT:  Godowski, Paul J.
;  APPLICANT:  Grimaldi, Christopher J.
;  APPLICANT:  Gurney, Austin L.
;  APPLICANT:  Hillan, Kenneth, J.
;  APPLICANT:  Kljavin, Ivar J.
;  APPLICANT:  Mather, Jennie P.
;  APPLICANT:  Pan, James
;  APPLICANT:  Paoni, Nicholas F.
;  APPLICANT:  Roy, Margaret Ann
;  APPLICANT:  Stewart, Timothy A.
;  APPLICANT:  Tumas, Daniel
;  APPLICANT:  Williams, P. Mickey
;  APPLICANT:  Wood, William, I.
;  TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
;  TITLE OF INVENTION:  Acids Encoding the Same
;  FILE REFERENCE: 10466-14
;  CURRENT APPLICATION NUMBER: US/09/909,088B
;  CURRENT FILING DATE:  2001-07-18
;  PRIOR APPLICATION NUMBER: PCT/US00/04414
;  PRIOR FILING DATE: 2000-02-22
;  PRIOR APPLICATION NUMBER: US 60/143,048
;  PRIOR FILING DATE: 1999-07-07
;  PRIOR APPLICATION NUMBER: US 60/145,698
;  PRIOR FILING DATE: 1999-07-26
;  PRIOR APPLICATION NUMBER: US 60/146,222
;  PRIOR FILING DATE: 1999-07-28
;  PRIOR APPLICATION NUMBER: PCT/US99/20594
;  PRIOR FILING DATE: 1999-09-08
;  PRIOR APPLICATION NUMBER: PCT/US99/20944
;  PRIOR FILING DATE: 1999-09-13
;  PRIOR APPLICATION NUMBER: PCT/US99/21090
;  PRIOR FILING DATE: 1999-09-15
;  PRIOR APPLICATION NUMBER: PCT/US99/21547
;  PRIOR FILING DATE: 1999-09-15
;  PRIOR APPLICATION NUMBER: PCT/US99/23089
;  PRIOR FILING DATE: 1999-10-05
;  PRIOR APPLICATION NUMBER: PCT/US99/28214
;  PRIOR FILING DATE: 1999-11-29
;  PRIOR APPLICATION NUMBER: PCT/US99/28313
;  PRIOR FILING DATE: 1999-11-30
;  PRIOR APPLICATION NUMBER: PCT/US99/28564
;  PRIOR FILING DATE: 1999-12-02
;  PRIOR APPLICATION NUMBER: PCT/US99/28565
;  PRIOR FILING DATE: 1999-12-02
;  PRIOR APPLICATION NUMBER: PCT/US99/30095
;  PRIOR FILING DATE: 1999-12-16
;  PRIOR APPLICATION NUMBER: PCT/US99/30911
;  PRIOR FILING DATE: 1999-12-20
;  PRIOR APPLICATION NUMBER: PCT/US99/30999
;  PRIOR FILING DATE: 1999-12-20
;  PRIOR APPLICATION NUMBER: PCT/US00/00219
;  PRIOR FILING DATE: 2000-01-05
;  NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 137
;   LENGTH: 382
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-909-088B-137
```

```
  Query Match             98.7%;  Score 1101;  DB 9;  Length 382;
  Best Local Similarity   96.3%;  Pred. No. 8.8e-103;
  Matches  206;  Conservative    0;  Mismatches    0;  Indels    8;  Gaps    1;

Qy          1 ATGRLLS--------GQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 52
              |||||||        ||||||||||||||||||||||||||||||||||||||||||||
Db         22 ATGRLLSASDLDLRGGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 81

Qy         53 ESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVD 112
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 ESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVD 141

Qy        113 EPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGE 172
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 EPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGE 201

Qy        173 ETELTTPVLPEETQEEDAKKTFKESREAALNLAY 206
              |||||||||||||||||||||||||||||||||
Db        202 ETELTTPVLPEETQEEDAKKTFKESREAALNLAY 235


RESULT 12
US-09-905-291A-137
; Sequence 137, Application US/09905291A
; Patent No. US20020160374A1
; GENERAL INFORMATION:
;  APPLICANT: Genentech, Inc.
;  APPLICANT:  Ashkenazi, Avi
;  APPLICANT:  Botstein, David
;  APPLICANT:  Desnoyers, Luc
;  APPLICANT:  Eaton, Dan L.
;  APPLICANT:  Ferrara, Napoleone
;  APPLICANT:  Filvaroff, Ellen
;  APPLICANT:  Fong, Sherman
;  APPLICANT:  Gao, Wei-Qiang
;  APPLICANT:  Gerber, Hanspeter
;  APPLICANT:  Gerritsen, Mary E.
;  APPLICANT:  Goddard, A.
;  APPLICANT:  Godowski, Paul J.
;  APPLICANT:  Grimaldi, Christopher J.
;  APPLICANT:  Gurney, Austin L.
;  APPLICANT:  Hillan, Kenneth, J.
;  APPLICANT:  Kljavin, Ivar J.
;  APPLICANT:  Mather, Jennie P.
;  APPLICANT:  Pan, James
;  APPLICANT:  Paoni, Nicholas F.
;  APPLICANT:  Roy, Margaret Ann
;  APPLICANT:  Stewart, Timothy A.
;  APPLICANT:  Tumas, Daniel
;  APPLICANT:  Williams, P. Mickey
;  APPLICANT:  Wood, William, I.
;  TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
;  TITLE OF INVENTION:  Acids Encoding the Same
;  FILE REFERENCE: 10466-14
;  CURRENT APPLICATION NUMBER: US/09/905,291A
;  CURRENT FILING DATE:  2001-07-12
;  PRIOR APPLICATION NUMBER: PCT/US00/04414
;  PRIOR FILING DATE: 2000-02-22
;  PRIOR APPLICATION NUMBER: US 60/143,048
;  PRIOR FILING DATE: 1999-07-07
;  PRIOR APPLICATION NUMBER: US 60/145,698
;  PRIOR FILING DATE: 1999-07-26
;  PRIOR APPLICATION NUMBER: US 60/146,222
;  PRIOR FILING DATE: 1999-07-28
;  PRIOR APPLICATION NUMBER: PCT/US99/20594
;  PRIOR FILING DATE: 1999-09-08
;  PRIOR APPLICATION NUMBER: PCT/US99/20944
;  PRIOR FILING DATE: 1999-09-13
;  PRIOR APPLICATION NUMBER: PCT/US99/21090
;  PRIOR FILING DATE: 1999-09-15
;  PRIOR APPLICATION NUMBER: PCT/US99/21547
;  PRIOR FILING DATE: 1999-09-15
;  PRIOR APPLICATION NUMBER: PCT/US99/23089
;  PRIOR FILING DATE: 1999-10-05
;  PRIOR APPLICATION NUMBER: PCT/US99/28214
;  PRIOR FILING DATE: 1999-11-29
;  PRIOR APPLICATION NUMBER: PCT/US99/28313
;  PRIOR FILING DATE: 1999-11-30
;  PRIOR APPLICATION NUMBER: PCT/US99/28564
;  PRIOR FILING DATE: 1999-12-02
;  PRIOR APPLICATION NUMBER: PCT/US99/28565
;  PRIOR FILING DATE: 1999-12-02
;  PRIOR APPLICATION NUMBER: PCT/US99/30095
;  PRIOR FILING DATE: 1999-12-16
;  PRIOR APPLICATION NUMBER: PCT/US99/30911
;  PRIOR FILING DATE: 1999-12-20
;  PRIOR APPLICATION NUMBER: PCT/US99/30999
;  PRIOR FILING DATE: 1999-12-20
;  PRIOR APPLICATION NUMBER: PCT/US00/00219
;  PRIOR FILING DATE: 2000-01-05
;  NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 137
;   LENGTH: 382
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-905-291A-137

  Query Match             98.7%;  Score 1101;  DB 9;  Length 382;
  Best Local Similarity   96.3%;  Pred. No. 8.8e-103;
  Matches  206;  Conservative    0;  Mismatches    0;  Indels    8;  Gaps    1;

Qy          1 ATGRLLS--------GQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 52
              |||||||        ||||||||||||||||||||||||||||||||||||||||||||
Db         22 ATGRLLSASDLDLRGGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 81

Qy         53 ESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVD 112
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 ESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVD 141

Qy        113 EPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGE 172
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 EPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGE 201

Qy        173 ETELTTPVLPEETQEEDAKKTFKESREAALNLAY 206
              |||||||||||||||||||||||||||||||||
Db        202 ETELTTPVLPEETQEEDAKKTFKESREAALNLAY 235


RESULT 13
US-09-902-853-137
; Sequence 137, Application US/09902853
; Publication No. US20020192659A1
; GENERAL INFORMATION:
;  APPLICANT: Genentech, Inc.
;  APPLICANT:  Ashkenazi, Avi
;  APPLICANT:  Botstein, David
;  APPLICANT:  Desnoyers, Luc
;  APPLICANT:  Eaton, Dan L.
;  APPLICANT:  Ferrara, Napoleone
;  APPLICANT:  Filvaroff, Ellen
;  APPLICANT:  Fong, Sherman
;  APPLICANT:  Gao, Wei-Qiang
;  APPLICANT:  Gerber, Hanspeter
;  APPLICANT:  Gerritsen, Mary E.
;  APPLICANT:  Goddard, A.
;  APPLICANT:  Godowski, Paul J.
;  APPLICANT:  Grimaldi, Christopher J.
;  APPLICANT:  Gurney, Austin L.
;  APPLICANT:  Hillan, Kenneth, J.
;  APPLICANT:  Kljavin, Ivar J.
;  APPLICANT:  Mather, Jennie P.
;  APPLICANT:  Pan, James
;  APPLICANT:  Paoni, Nicholas F.
;  APPLICANT:  Roy, Margaret Ann
;  APPLICANT:  Stewart, Timothy A.
```

```
;  APPLICANT:  Tumas, Daniel
;  APPLICANT:  Williams, P. Mickey
;  APPLICANT:  Wood, William, I.
;  TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
;  TITLE OF INVENTION:  Acids Encoding the Same
;  FILE REFERENCE: 10466-14
;  CURRENT APPLICATION NUMBER: US/09/902,853
;  CURRENT FILING DATE:  2001-07-10
;  PRIOR APPLICATION NUMBER: US/09/665,350
;  PRIOR FILING DATE: 2000-09-18
;  PRIOR APPLICATION NUMBER: US 60/143,048
;  PRIOR FILING DATE: 1999-07-07
;  PRIOR APPLICATION NUMBER: US 60/145,698
;  PRIOR FILING DATE: 1999-07-26
;  PRIOR APPLICATION NUMBER: US 60/146,222
;  PRIOR FILING DATE: 1999-07-28
;  PRIOR APPLICATION NUMBER: PCT/US99/20594
;  PRIOR FILING DATE: 1999-09-08
;  PRIOR APPLICATION NUMBER: PCT/US99/20944
;  PRIOR FILING DATE: 1999-09-13
;  PRIOR APPLICATION NUMBER: PCT/US99/21090
;  PRIOR FILING DATE: 1999-09-15
;  PRIOR APPLICATION NUMBER: PCT/US99/21547
;  PRIOR FILING DATE: 1999-09-15
;  PRIOR APPLICATION NUMBER: PCT/US99/23089
;  PRIOR FILING DATE: 1999-10-05
;  PRIOR APPLICATION NUMBER: PCT/US99/28214
;  PRIOR FILING DATE: 1999-11-29
;  PRIOR APPLICATION NUMBER: PCT/US99/28313
;  PRIOR FILING DATE: 1999-11-30
;  PRIOR APPLICATION NUMBER: PCT/US99/28564
;  PRIOR FILING DATE: 1999-12-02
;  PRIOR APPLICATION NUMBER: PCT/US99/28565
;  PRIOR FILING DATE: 1999-12-02
;  PRIOR APPLICATION NUMBER: PCT/US99/30095
;  PRIOR FILING DATE: 1999-12-16
;  PRIOR APPLICATION NUMBER: PCT/US99/30911
;  PRIOR FILING DATE: 1999-12-20
;  PRIOR APPLICATION NUMBER: PCT/US99/30999
;  PRIOR FILING DATE: 1999-12-20
;  PRIOR APPLICATION NUMBER: PCT/US00/00219
;  PRIOR FILING DATE: 2000-01-05
;  NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 137
;    LENGTH: 382
;    TYPE: PRT
;    ORGANISM: Homo Sapien
US-09-902-853-137

  Query Match             98.7%;  Score 1101;  DB 9;  Length 382;
  Best Local Similarity   96.3%;  Pred. No. 8.8e-103;
  Matches  206;  Conservative    0;  Mismatches    0;  Indels    8;  Gaps    1;

Qy          1 ATGRLLS--------GQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 52
              |||||||        ||||||||||||||||||||||||||||||||||||||||||||
Db         22 ATGRLLSASDLDLRGGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 81

Qy         53 ESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVD 112
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 ESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVD 141

Qy        113 EPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGE 172
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 EPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGE 201

Qy        173 ETELTTPVLPEETQEEDAKKTFKESREAALNLAY 206
              |||||||||||||||||||||||||||||||||
Db        202 ETELTTPVLPEETQEEDAKKTFKESREAALNLAY 235


RESULT 14
US-09-907-824-137
; Sequence 137, Application US/09907824
; Publication No. US20020197671A1
; GENERAL INFORMATION:
;  APPLICANT: Genentech, Inc.
;  APPLICANT:  Ashkenazi, Avi
;  APPLICANT:  Botstein, David
;  APPLICANT:  Desnoyers, Luc
;  APPLICANT:  Eaton, Dan L.
;  APPLICANT:  Ferrara, Napoleone
;  APPLICANT:  Filvaroff, Ellen
;  APPLICANT:  Fong, Sherman
;  APPLICANT:  Gao, Wei-Qiang
;  APPLICANT:  Gerber, Hanspeter
;  APPLICANT:  Gerritsen, Mary E.
;  APPLICANT:  Goddard, A.
;  APPLICANT:  Godowski, Paul J.
;  APPLICANT:  Grimaldi, Christopher J.
;  APPLICANT:  Gurney, Austin L.
;  APPLICANT:  Hillan, Kenneth, J.
;  APPLICANT:  Kljavin, Ivar J.
;  APPLICANT:  Mather, Jennie P.
;  APPLICANT:  Pan, James
;  APPLICANT:  Paoni, Nicholas F.
;  APPLICANT:  Roy, Margaret Ann
;  APPLICANT:  Stewart, Timothy A.
;  APPLICANT:  Tumas, Daniel
;  APPLICANT:  Williams, P. Mickey
;  APPLICANT:  Wood, William, I.
;  TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
;  TITLE OF INVENTION:  Acids Encoding the Same
;  FILE REFERENCE: 10466-14
;  CURRENT APPLICATION NUMBER: US/09/907,824
;  CURRENT FILING DATE:  2001-07-17
;  PRIOR APPLICATION NUMBER: 09/665,350
;  PRIOR FILING DATE: 2000-09-18
;  PRIOR APPLICATION NUMBER: PCT/US00/04414
;  PRIOR FILING DATE: 2000-02-22
;  PRIOR APPLICATION NUMBER: US 60/143,048
;  PRIOR FILING DATE: 1999-07-07
;  PRIOR APPLICATION NUMBER: US 60/145,698
;  PRIOR FILING DATE: 1999-07-26
;  PRIOR APPLICATION NUMBER: US 60/146,222
;  PRIOR FILING DATE: 1999-07-28
;  PRIOR APPLICATION NUMBER: PCT/US99/20594
;  PRIOR FILING DATE: 1999-09-08
;  PRIOR APPLICATION NUMBER: PCT/US99/20944
;  PRIOR FILING DATE: 1999-09-13
;  PRIOR APPLICATION NUMBER: PCT/US99/21090
;  PRIOR FILING DATE: 1999-09-15
;  PRIOR APPLICATION NUMBER: PCT/US99/21547
;  PRIOR FILING DATE: 1999-09-15
;  PRIOR APPLICATION NUMBER: PCT/US99/23089
;  PRIOR FILING DATE: 1999-10-05
;  PRIOR APPLICATION NUMBER: PCT/US99/28214
;  PRIOR FILING DATE: 1999-11-29
;  PRIOR APPLICATION NUMBER: PCT/US99/28313
;  PRIOR FILING DATE: 1999-11-30
;  PRIOR APPLICATION NUMBER: PCT/US99/28564
;  PRIOR FILING DATE: 1999-12-02
;  PRIOR APPLICATION NUMBER: PCT/US99/28565
;  PRIOR FILING DATE: 1999-12-02
;  PRIOR APPLICATION NUMBER: PCT/US99/30095
;  PRIOR FILING DATE: 1999-12-16
;  PRIOR APPLICATION NUMBER: PCT/US99/30911
;  PRIOR FILING DATE: 1999-12-20
;  PRIOR APPLICATION NUMBER: PCT/US99/30999
;  PRIOR FILING DATE: 1999-12-20
;  PRIOR APPLICATION NUMBER: PCT/US00/00219
;  PRIOR FILING DATE: 2000-01-05
;  NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 137
;    LENGTH: 382
;    TYPE: PRT
;    ORGANISM: Homo Sapien
```

US-09-907-824-137

Query Match 98.7%; Score 1101; DB 9; Length 382;
Best Local Similarity 96.3%; Pred. No. 8.8e-103;
Matches 206; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

```
Qy          1 ATGRLLS--------GQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 52
              |||||||        ||||||||||||||||||||||||||||||||||||||||||||
Db         22 ATGRLLSASDLDLRGGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 81

Qy         53 ESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVD 112
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 ESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVD 141

Qy        113 EPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGE 172
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 EPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGE 201

Qy        173 ETELTTPVLPEETQEEDAKKTFKESREAALNLAY 206
              |||||||||||||||||||||||||||||||||
Db        202 ETELTTPVLPEETQEEDAKKTFKESREAALNLAY 235
```

RESULT 15
US-09-907-841-137
; Sequence 137, Application US/09907841
; Publication No. US20020198366A1
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/907,841
; CURRENT FILING DATE: 2001-11-20
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 137
; LENGTH: 382
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-907-841-137

Query Match 98.7%; Score 1101; DB 9; Length 382;
Best Local Similarity 96.3%; Pred. No. 8.8e-103;
Matches 206; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

```
Qy          1 ATGRLLS--------GQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 52
              |||||||        ||||||||||||||||||||||||||||||||||||||||||||
Db         22 ATGRLLSASDLDLRGGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 81

Qy         53 ESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVD 112
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 ESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVD 141

Qy        113 EPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGE 172
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 EPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGE 201

Qy        173 ETELTTPVLPEETQEEDAKKTFKESREAALNLAY 206
              |||||||||||||||||||||||||||||||||
Db        202 ETELTTPVLPEETQEEDAKKTFKESREAALNLAY 235
```

RESULT 16
US-09-904-011-137
; Sequence 137, Application US/09904011
; Publication No. US20030003530A1
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/904,011
; CURRENT FILING DATE: 2001-07-11
; PRIOR APPLICATION NUMBER: 09/665,350
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22

```
;  PRIOR APPLICATION NUMBER: US 60/143,048
;  PRIOR FILING DATE: 1999-07-07
;  PRIOR APPLICATION NUMBER: US 60/145,698
;  PRIOR FILING DATE: 1999-07-26
;  PRIOR APPLICATION NUMBER: US 60/146,222
;  PRIOR FILING DATE: 1999-07-28
;  PRIOR APPLICATION NUMBER: PCT/US99/20594
;  PRIOR FILING DATE: 1999-09-08
;  PRIOR APPLICATION NUMBER: PCT/US99/20944
;  PRIOR FILING DATE: 1999-09-13
;  PRIOR APPLICATION NUMBER: PCT/US99/21090
;  PRIOR FILING DATE: 1999-09-15
;  PRIOR APPLICATION NUMBER: PCT/US99/21547
;  PRIOR FILING DATE: 1999-09-15
;  PRIOR APPLICATION NUMBER: PCT/US99/23089
;  PRIOR FILING DATE: 1999-10-05
;  PRIOR APPLICATION NUMBER: PCT/US99/28214
;  PRIOR FILING DATE: 1999-11-29
;  PRIOR APPLICATION NUMBER: PCT/US99/28313
;  PRIOR FILING DATE: 1999-11-30
;  PRIOR APPLICATION NUMBER: PCT/US99/28564
;  PRIOR FILING DATE: 1999-12-02
;  PRIOR APPLICATION NUMBER: PCT/US99/28565
;  PRIOR FILING DATE: 1999-12-02
;  PRIOR APPLICATION NUMBER: PCT/US99/30095
;  PRIOR FILING DATE: 1999-12-16
;  PRIOR APPLICATION NUMBER: PCT/US99/30911
;  PRIOR FILING DATE: 1999-12-20
;  PRIOR APPLICATION NUMBER: PCT/US99/30999
;  PRIOR FILING DATE: 1999-12-20
;  PRIOR APPLICATION NUMBER: PCT/US00/00219
;  PRIOR FILING DATE: 2000-01-05
;  NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 137
;   LENGTH: 382
;   TYPE: PRT
;   ORGANISM: Homo Sapien
US-09-904-011-137

  Query Match             98.7%;  Score 1101;  DB 10;  Length 382;
  Best Local Similarity   96.3%;  Pred. No. 8.8e-103;
  Matches  206;  Conservative    0;  Mismatches    0;  Indels    8;  Gaps    1;

Qy          1 ATGRLLS--------GQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 52
              |||||||        ||||||||||||||||||||||||||||||||||||||||||||
Db         22 ATGRLLSASDLDLRGGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 81

Qy         53 ESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVD 112
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 ESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVD 141

Qy        113 EPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGE 172
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 EPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGE 201

Qy        173 ETELTTPVLPEETQEEDAKKTFKESREAALNLAY 206
              |||||||||||||||||||||||||||||||||
Db        202 ETELTTPVLPEETQEEDAKKTFKESREAALNLAY 235


RESULT 17
US-09-906-742-137
; Sequence 137, Application US/09906742
; Publication No. US20030023054A1
; GENERAL INFORMATION:
;  APPLICANT: Genentech, Inc.
;  APPLICANT:  Ashkenazi, Avi
;  APPLICANT:  Botstein, David
;  APPLICANT:  Desnoyers, Luc
;  APPLICANT:  Eaton, Dan L.
;  APPLICANT:  Ferrara, Napoleone
;  APPLICANT:  Filvaroff, Ellen
;  APPLICANT:  Fong, Sherman
;  APPLICANT:  Gao, Wei-Qiang
;  APPLICANT:  Gerber, Hanspeter
;  APPLICANT:  Gerritsen, Mary E.
;  APPLICANT:  Goddard, A.
;  APPLICANT:  Godowski, Paul J.
;  APPLICANT:  Grimaldi, Christopher J.
;  APPLICANT:  Gurney, Austin L.
;  APPLICANT:  Hillan, Kenneth, J.
;  APPLICANT:  Kljavin, Ivar J.
;  APPLICANT:  Mather, Jennie P.
;  APPLICANT:  Pan, James
;  APPLICANT:  Paoni, Nicholas F.
;  APPLICANT:  Roy, Margaret Ann
;  APPLICANT:  Stewart, Timothy A.
;  APPLICANT:  Tumas, Daniel
;  APPLICANT:  Williams, P. Mickey
;  APPLICANT:  Wood, William, I.
;  TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
;  TITLE OF INVENTION:  Acids Encoding the Same
;  FILE REFERENCE: 10466-14
;  CURRENT APPLICATION NUMBER: US/09/906,742
;  CURRENT FILING DATE:  2001-07-16
;  PRIOR APPLICATION NUMBER: 09/665,350
;  PRIOR FILING DATE: 2000-09-18
;  PRIOR APPLICATION NUMBER: PCT/US00/04414
;  PRIOR FILING DATE: 2000-02-22
;  PRIOR APPLICATION NUMBER: US 60/143,048
;  PRIOR FILING DATE: 1999-07-07
;  PRIOR APPLICATION NUMBER: US 60/145,698
;  PRIOR FILING DATE: 1999-07-26
;  PRIOR APPLICATION NUMBER: US 60/146,222
;  PRIOR FILING DATE: 1999-07-28
;  PRIOR APPLICATION NUMBER: PCT/US99/20594
;  PRIOR FILING DATE: 1999-09-08
;  PRIOR APPLICATION NUMBER: PCT/US99/20944
;  PRIOR FILING DATE: 1999-09-13
;  PRIOR APPLICATION NUMBER: PCT/US99/21090
;  PRIOR FILING DATE: 1999-09-15
;  PRIOR APPLICATION NUMBER: PCT/US99/21547
;  PRIOR FILING DATE: 1999-09-15
;  PRIOR APPLICATION NUMBER: PCT/US99/23089
;  PRIOR FILING DATE: 1999-10-05
;  PRIOR APPLICATION NUMBER: PCT/US99/28214
;  PRIOR FILING DATE: 1999-11-29
;  PRIOR APPLICATION NUMBER: PCT/US99/28313
;  PRIOR FILING DATE: 1999-11-30
;  PRIOR APPLICATION NUMBER: PCT/US99/28564
;  PRIOR FILING DATE: 1999-12-02
;  PRIOR APPLICATION NUMBER: PCT/US99/28565
;  PRIOR FILING DATE: 1999-12-02
;  PRIOR APPLICATION NUMBER: PCT/US99/30095
;  PRIOR FILING DATE: 1999-12-16
;  PRIOR APPLICATION NUMBER: PCT/US99/30911
;  PRIOR FILING DATE: 1999-12-20
;  PRIOR APPLICATION NUMBER: PCT/US99/30999
;  PRIOR FILING DATE: 1999-12-20
;  PRIOR APPLICATION NUMBER: PCT/US00/00219
;  PRIOR FILING DATE: 2000-01-05
;  NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 137
;   LENGTH: 382
;   TYPE: PRT
;   ORGANISM: Homo Sapien
US-09-906-742-137

  Query Match             98.7%;  Score 1101;  DB 10;  Length 382;
  Best Local Similarity   96.3%;  Pred. No. 8.8e-103;
  Matches  206;  Conservative    0;  Mismatches    0;  Indels    8;  Gaps    1;

Qy          1 ATGRLLS--------GQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 52
              |||||||        ||||||||||||||||||||||||||||||||||||||||||||
Db         22 ATGRLLSASDLDLRGGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 81
```

```
Qy          53 ESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVD 112
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          82 ESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVD 141

Qy         113 EPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGE 172
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         142 EPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGE 201

Qy         173 ETELTTPVLPEETQEEDAKKTFKESREAALNLAY 206
               |||||||||||||||||||||||||||||||||
Db         202 ETELTTPVLPEETQEEDAKKTFKESREAALNLAY 235


RESULT 18
US-09-906-838-137
; Sequence 137, Application US/09906838
; Publication No. US20030027143A1
; GENERAL INFORMATION:
;  APPLICANT: Genentech, Inc.
;  APPLICANT:  Ashkenazi, Avi
;  APPLICANT:  Botstein, David
;  APPLICANT:  Desnoyers, Luc
;  APPLICANT:  Eaton, Dan L.
;  APPLICANT:  Ferrara, Napoleone
;  APPLICANT:  Filvaroff, Ellen
;  APPLICANT:  Fong, Sherman
;  APPLICANT:  Gao, Wei-Qiang
;  APPLICANT:  Gerber, Hanspeter
;  APPLICANT:  Gerritsen, Mary E.
;  APPLICANT:  Goddard, A.
;  APPLICANT:  Godowski, Paul J.
;  APPLICANT:  Grimaldi, Christopher J.
;  APPLICANT:  Gurney, Austin L.
;  APPLICANT:  Hillan, Kenneth, J.
;  APPLICANT:  Kljavin, Ivar J.
;  APPLICANT:  Mather, Jennie P.
;  APPLICANT:  Pan, James
;  APPLICANT:  Paoni, Nicholas F.
;  APPLICANT:  Roy, Margaret Ann
;  APPLICANT:  Stewart, Timothy A.
;  APPLICANT:  Tumas, Daniel
;  APPLICANT:  Williams, P. Mickey
;  APPLICANT:  Wood, William, I.
;  TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
;  TITLE OF INVENTION:  Acids Encoding the Same
;  FILE REFERENCE: 10466-14
;  CURRENT APPLICATION NUMBER: US/09/906,838
;  CURRENT FILING DATE:  2001-07-16
;  PRIOR APPLICATION NUMBER: 09/665,350
;  PRIOR FILING DATE: 2000-09-18
;  PRIOR APPLICATION NUMBER: PCT/US00/04414
;  PRIOR FILING DATE: 2000-02-22
;  PRIOR APPLICATION NUMBER: US 60/143,048
;  PRIOR FILING DATE: 1999-07-07
;  PRIOR APPLICATION NUMBER: US 60/145,698
;  PRIOR FILING DATE: 1999-07-26
;  PRIOR APPLICATION NUMBER: US 60/146,222
;  PRIOR FILING DATE: 1999-07-28
;  PRIOR APPLICATION NUMBER: PCT/US99/20594
;  PRIOR FILING DATE: 1999-09-08
;  PRIOR APPLICATION NUMBER: PCT/US99/20944
;  PRIOR FILING DATE: 1999-09-13
;  PRIOR APPLICATION NUMBER: PCT/US99/21090
;  PRIOR FILING DATE: 1999-09-15
;  PRIOR APPLICATION NUMBER: PCT/US99/21547
;  PRIOR FILING DATE: 1999-09-15
;  PRIOR APPLICATION NUMBER: PCT/US99/23089
;  PRIOR FILING DATE: 1999-10-05
;  PRIOR APPLICATION NUMBER: PCT/US99/28214
;  PRIOR FILING DATE: 1999-11-29
;  PRIOR APPLICATION NUMBER: PCT/US99/28313
;  PRIOR FILING DATE: 1999-11-30
;  PRIOR APPLICATION NUMBER: PCT/US99/28564
;  PRIOR FILING DATE: 1999-12-02
;  PRIOR APPLICATION NUMBER: PCT/US99/28565
;  PRIOR FILING DATE: 1999-12-02
;  PRIOR APPLICATION NUMBER: PCT/US99/30095
;  PRIOR FILING DATE: 1999-12-16
;  PRIOR APPLICATION NUMBER: PCT/US99/30911
;  PRIOR FILING DATE: 1999-12-20
;  PRIOR APPLICATION NUMBER: PCT/US99/30999
;  PRIOR FILING DATE: 1999-12-20
;  PRIOR APPLICATION NUMBER: PCT/US00/00219
;  PRIOR FILING DATE: 2000-01-05
;  NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 137
;   LENGTH: 382
;   TYPE: PRT
;   ORGANISM: Homo Sapien
US-09-906-838-137

  Query Match             98.7%;  Score 1101;  DB 10;  Length 382;
  Best Local Similarity   96.3%;  Pred. No. 8.8e-103;
  Matches  206;  Conservative    0;  Mismatches    0;  Indels    8;  Gaps    1;

Qy           1 ATGRLLS--------GQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 52
               |||||||        ||||||||||||||||||||||||||||||||||||||||||||
Db          22 ATGRLLSASDLDLRGGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 81

Qy          53 ESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVD 112
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          82 ESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVD 141

Qy         113 EPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGE 172
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         142 EPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGE 201

Qy         173 ETELTTPVLPEETQEEDAKKTFKESREAALNLAY 206
               |||||||||||||||||||||||||||||||||
Db         202 ETELTTPVLPEETQEEDAKKTFKESREAALNLAY 235


RESULT 19
US-09-907-613-137
; Sequence 137, Application US/09907613
; Publication No. US20030027145A1
; GENERAL INFORMATION:
;  APPLICANT: Genentech, Inc.
;  APPLICANT:  Ashkenazi, Avi
;  APPLICANT:  Botstein, David
;  APPLICANT:  Desnoyers, Luc
;  APPLICANT:  Eaton, Dan L.
;  APPLICANT:  Ferrara, Napoleone
;  APPLICANT:  Filvaroff, Ellen
;  APPLICANT:  Fong, Sherman
;  APPLICANT:  Gao, Wei-Qiang
;  APPLICANT:  Gerber, Hanspeter
;  APPLICANT:  Gerritsen, Mary E.
;  APPLICANT:  Goddard, A.
;  APPLICANT:  Godowski, Paul J.
;  APPLICANT:  Grimaldi, Christopher J.
;  APPLICANT:  Gurney, Austin L.
;  APPLICANT:  Hillan, Kenneth, J.
;  APPLICANT:  Kljavin, Ivar J.
;  APPLICANT:  Mather, Jennie P.
;  APPLICANT:  Pan, James
;  APPLICANT:  Paoni, Nicholas F.
;  APPLICANT:  Roy, Margaret Ann
;  APPLICANT:  Stewart, Timothy A.
;  APPLICANT:  Tumas, Daniel
;  APPLICANT:  Williams, P. Mickey
;  APPLICANT:  Wood, William, I.
;  TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
;  TITLE OF INVENTION:  Acids Encoding the Same
;  FILE REFERENCE: 10466-14
;  CURRENT APPLICATION NUMBER: US/09/907,613
```

```
;  CURRENT FILING DATE:  2001-07-17
;  PRIOR APPLICATION NUMBER: PCT/US00/04414
;  PRIOR FILING DATE: 2000-02-22
;  PRIOR APPLICATION NUMBER: US 60/143,048
;  PRIOR FILING DATE: 1999-07-07
;  PRIOR APPLICATION NUMBER: US 60/145,698
;  PRIOR FILING DATE: 1999-07-26
;  PRIOR APPLICATION NUMBER: US 60/146,222
;  PRIOR FILING DATE: 1999-07-28
;  PRIOR APPLICATION NUMBER: PCT/US99/20594
;  PRIOR FILING DATE: 1999-09-08
;  PRIOR APPLICATION NUMBER: PCT/US99/20944
;  PRIOR FILING DATE: 1999-09-13
;  PRIOR APPLICATION NUMBER: PCT/US99/21090
;  PRIOR FILING DATE: 1999-09-15
;  PRIOR APPLICATION NUMBER: PCT/US99/21547
;  PRIOR FILING DATE: 1999-09-15
;  PRIOR APPLICATION NUMBER: PCT/US99/23089
;  PRIOR FILING DATE: 1999-10-05
;  PRIOR APPLICATION NUMBER: PCT/US99/28214
;  PRIOR FILING DATE: 1999-11-29
;  PRIOR APPLICATION NUMBER: PCT/US99/28313
;  PRIOR FILING DATE: 1999-11-30
;  PRIOR APPLICATION NUMBER: PCT/US99/28564
;  PRIOR FILING DATE: 1999-12-02
;  PRIOR APPLICATION NUMBER: PCT/US99/28565
;  PRIOR FILING DATE: 1999-12-02
;  PRIOR APPLICATION NUMBER: PCT/US99/30095
;  PRIOR FILING DATE: 1999-12-16
;  PRIOR APPLICATION NUMBER: PCT/US99/30911
;  PRIOR FILING DATE: 1999-12-20
;  PRIOR APPLICATION NUMBER: PCT/US99/30999
;  PRIOR FILING DATE: 1999-12-20
;  PRIOR APPLICATION NUMBER: PCT/US00/00219
;  PRIOR FILING DATE: 2000-01-05
;  NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 137
;   LENGTH: 382
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-907-613-137

  Query Match             98.7%;  Score 1101;  DB 10;  Length 382;
  Best Local Similarity   96.3%;  Pred. No. 8.8e-103;
  Matches  206;  Conservative    0;  Mismatches    0;  Indels    8;  Gaps    1;

Qy          1 ATGRLLS--------GQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 52
              |||||||        ||||||||||||||||||||||||||||||||||||||||||||
Db         22 ATGRLLSASDLDLRGGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 81

Qy         53 ESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVD 112
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 ESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVD 141

Qy        113 EPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGE 172
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 EPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGE 201

Qy        173 ETELTTPVLPEETQEEDAKKTFKESREAALNLAY 206
              |||||||||||||||||||||||||||||||||
Db        202 ETELTTPVLPEETQEEDAKKTFKESREAALNLAY 235


RESULT 20
US-09-907-942-137
; Sequence 137, Application US/09907942
; Publication No. US20030027146A1
; GENERAL INFORMATION:
;  APPLICANT: Genentech, Inc.
;  APPLICANT:  Ashkenazi, Avi
;  APPLICANT:  Botstein, David
;  APPLICANT:  Desnoyers, Luc
;  APPLICANT:  Eaton, Dan L.
;  APPLICANT:  Ferrara, Napoleone
;  APPLICANT:  Filvaroff, Ellen
;  APPLICANT:  Fong, Sherman
;  APPLICANT:  Gao, Wei-Qiang
;  APPLICANT:  Gerber, Hanspeter
;  APPLICANT:  Gerritsen, Mary E.
;  APPLICANT:  Goddard, A.
;  APPLICANT:  Godowski, Paul J.
;  APPLICANT:  Grimaldi, Christopher J.
;  APPLICANT:  Gurney, Austin L.
;  APPLICANT:  Hillan, Kenneth, J.
;  APPLICANT:  Kljavin, Ivar J.
;  APPLICANT:  Mather, Jennie P.
;  APPLICANT:  Pan, James
;  APPLICANT:  Paoni, Nicholas F.
;  APPLICANT:  Roy, Margaret Ann
;  APPLICANT:  Stewart, Timothy A.
;  APPLICANT:  Tumas, Daniel
;  APPLICANT:  Williams, P. Mickey
;  APPLICANT:  Wood, William, I.
;  TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
;  TITLE OF INVENTION:  Acids Encoding the Same
;  FILE REFERENCE: 10466-14
;  CURRENT APPLICATION NUMBER: US/09/907,942
;  CURRENT FILING DATE:  2002-01-22
;  PRIOR APPLICATION NUMBER: PCT/US00/04414
;  PRIOR FILING DATE: 2000-02-22
;  PRIOR APPLICATION NUMBER: US 60/143,048
;  PRIOR FILING DATE: 1999-07-07
;  PRIOR APPLICATION NUMBER: US 60/145,698
;  PRIOR FILING DATE: 1999-07-26
;  PRIOR APPLICATION NUMBER: US 60/146,222
;  PRIOR FILING DATE: 1999-07-28
;  PRIOR APPLICATION NUMBER: PCT/US99/20594
;  PRIOR FILING DATE: 1999-09-08
;  PRIOR APPLICATION NUMBER: PCT/US99/20944
;  PRIOR FILING DATE: 1999-09-13
;  PRIOR APPLICATION NUMBER: PCT/US99/21090
;  PRIOR FILING DATE: 1999-09-15
;  PRIOR APPLICATION NUMBER: PCT/US99/21547
;  PRIOR FILING DATE: 1999-09-15
;  PRIOR APPLICATION NUMBER: PCT/US99/23089
;  PRIOR FILING DATE: 1999-10-05
;  PRIOR APPLICATION NUMBER: PCT/US99/28214
;  PRIOR FILING DATE: 1999-11-29
;  PRIOR APPLICATION NUMBER: PCT/US99/28313
;  PRIOR FILING DATE: 1999-11-30
;  PRIOR APPLICATION NUMBER: PCT/US99/28564
;  PRIOR FILING DATE: 1999-12-02
;  PRIOR APPLICATION NUMBER: PCT/US99/28565
;  PRIOR FILING DATE: 1999-12-02
;  PRIOR APPLICATION NUMBER: PCT/US99/30095
;  PRIOR FILING DATE: 1999-12-16
;  PRIOR APPLICATION NUMBER: PCT/US99/30911
;  PRIOR FILING DATE: 1999-12-20
;  PRIOR APPLICATION NUMBER: PCT/US99/30999
;  PRIOR FILING DATE: 1999-12-20
;  PRIOR APPLICATION NUMBER: PCT/US00/00219
;  PRIOR FILING DATE: 2000-01-05
;  NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 137
;   LENGTH: 382
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-907-942-137

  Query Match             98.7%;  Score 1101;  DB 10;  Length 382;
  Best Local Similarity   96.3%;  Pred. No. 8.8e-103;
  Matches  206;  Conservative    0;  Mismatches    0;  Indels    8;  Gaps    1;

Qy          1 ATGRLLS--------GQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 52
              |||||||        ||||||||||||||||||||||||||||||||||||||||||||
Db         22 ATGRLLSASDLDLRGGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 81
```

```
Qy          53 ESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVD 112
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          82 ESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVD 141

Qy         113 EPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGE 172
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         142 EPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGE 201

Qy         173 ETELTTPVLPEETQEEDAKKTFKESREAALNLAY 206
               |||||||||||||||||||||||||||||||||
Db         202 ETELTTPVLPEETQEEDAKKTFKESREAALNLAY 235


RESULT 21
US-09-904-859-137
; Sequence 137, Application US/09904859
; Publication No. US20030036060A1
; GENERAL INFORMATION:
;  APPLICANT: Genentech, Inc.
;  APPLICANT:  Ashkenazi, Avi
;  APPLICANT:  Botstein, David
;  APPLICANT:  Desnoyers, Luc
;  APPLICANT:  Eaton, Dan L.
;  APPLICANT:  Ferrara, Napoleone
;  APPLICANT:  Filvaroff, Ellen
;  APPLICANT:  Fong, Sherman
;  APPLICANT:  Gao, Wei-Qiang
;  APPLICANT:  Gerber, Hanspeter
;  APPLICANT:  Gerritsen, Mary E.
;  APPLICANT:  Goddard, A.
;  APPLICANT:  Godowski, Paul J.
;  APPLICANT:  Grimaldi, Christopher J.
;  APPLICANT:  Gurney, Austin L.
;  APPLICANT:  Hillan, Kenneth, J.
;  APPLICANT:  Kljavin, Ivar J.
;  APPLICANT:  Mather, Jennie P.
;  APPLICANT:  Pan, James
;  APPLICANT:  Paoni, Nicholas F.
;  APPLICANT:  Roy, Margaret Ann
;  APPLICANT:  Stewart, Timothy A.
;  APPLICANT:  Tumas, Daniel
;  APPLICANT:  Williams, P. Mickey
;  APPLICANT:  Wood, William, I.
;  TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
;  TITLE OF INVENTION:  Acids Encoding the Same
;  FILE REFERENCE: 10466-14
;  CURRENT APPLICATION NUMBER: US/09/904,859
;  CURRENT FILING DATE:  2001-07-12
;  PRIOR APPLICATION NUMBER: 09/665,350
;  PRIOR FILING DATE: 2000-09-18
;  PRIOR APPLICATION NUMBER: PCT/US00/04414
;  PRIOR FILING DATE: 2000-02-22
;  PRIOR APPLICATION NUMBER: US 60/143,048
;  PRIOR FILING DATE: 1999-07-07
;  PRIOR APPLICATION NUMBER: US 60/145,698
;  PRIOR FILING DATE: 1999-07-26
;  PRIOR APPLICATION NUMBER: US 60/146,222
;  PRIOR FILING DATE: 1999-07-28
;  PRIOR APPLICATION NUMBER: PCT/US99/20594
;  PRIOR FILING DATE: 1999-09-08
;  PRIOR APPLICATION NUMBER: PCT/US99/20944
;  PRIOR FILING DATE: 1999-09-13
;  PRIOR APPLICATION NUMBER: PCT/US99/21090
;  PRIOR FILING DATE: 1999-09-15
;  PRIOR APPLICATION NUMBER: PCT/US99/21547
;  PRIOR FILING DATE: 1999-09-15
;  PRIOR APPLICATION NUMBER: PCT/US99/23089
;  PRIOR FILING DATE: 1999-10-05
;  PRIOR APPLICATION NUMBER: PCT/US99/28214
;  PRIOR FILING DATE: 1999-11-29
;  PRIOR APPLICATION NUMBER: PCT/US99/28313
;  PRIOR FILING DATE: 1999-11-30
;  PRIOR APPLICATION NUMBER: PCT/US99/28564
;  PRIOR FILING DATE: 1999-12-02
;  PRIOR APPLICATION NUMBER: PCT/US99/28565
;  PRIOR FILING DATE: 1999-12-02
;  PRIOR APPLICATION NUMBER: PCT/US99/30095
;  PRIOR FILING DATE: 1999-12-16
;  PRIOR APPLICATION NUMBER: PCT/US99/30911
;  PRIOR FILING DATE: 1999-12-20
;  PRIOR APPLICATION NUMBER: PCT/US99/30999
;  PRIOR FILING DATE: 1999-12-20
;  PRIOR APPLICATION NUMBER: PCT/US00/00219
;  PRIOR FILING DATE: 2000-01-05
;  NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 137
;   LENGTH: 382
;   TYPE: PRT
;   ORGANISM: Homo Sapien
US-09-904-859-137

  Query Match             98.7%;  Score 1101;  DB 10;  Length 382;
  Best Local Similarity   96.3%;  Pred. No. 8.8e-103;
  Matches  206;  Conservative    0;  Mismatches    0;  Indels    8;  Gaps    1;

Qy           1 ATGRLLS--------GQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 52
                 |||||||        ||||||||||||||||||||||||||||||||||||||||||||
Db          22 ATGRLLSASDLDLRGGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 81

Qy          53 ESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVD 112
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          82 ESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVD 141

Qy         113 EPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGE 172
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         142 EPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGE 201

Qy         173 ETELTTPVLPEETQEEDAKKTFKESREAALNLAY 206
               |||||||||||||||||||||||||||||||||
Db         202 ETELTTPVLPEETQEEDAKKTFKESREAALNLAY 235


RESULT 22
US-09-909-204-137
; Sequence 137, Application US/09909204
; Publication No. US20030036061A1
; GENERAL INFORMATION:
;  APPLICANT: Genentech, Inc.
;  APPLICANT:  Ashkenazi, Avi
;  APPLICANT:  Botstein, David
;  APPLICANT:  Desnoyers, Luc
;  APPLICANT:  Eaton, Dan L.
;  APPLICANT:  Ferrara, Napoleone
;  APPLICANT:  Filvaroff, Ellen
;  APPLICANT:  Fong, Sherman
;  APPLICANT:  Gao, Wei-Qiang
;  APPLICANT:  Gerber, Hanspeter
;  APPLICANT:  Gerritsen, Mary E.
;  APPLICANT:  Goddard, A.
;  APPLICANT:  Godowski, Paul J.
;  APPLICANT:  Grimaldi, Christopher J.
;  APPLICANT:  Gurney, Austin L.
;  APPLICANT:  Hillan, Kenneth, J.
;  APPLICANT:  Kljavin, Ivar J.
;  APPLICANT:  Mather, Jennie P.
;  APPLICANT:  Pan, James
;  APPLICANT:  Paoni, Nicholas F.
;  APPLICANT:  Roy, Margaret Ann
;  APPLICANT:  Stewart, Timothy A.
;  APPLICANT:  Tumas, Daniel
;  APPLICANT:  Williams, P. Mickey
;  APPLICANT:  Wood, William, I.
;  TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
;  TITLE OF INVENTION:  Acids Encoding the Same
;  FILE REFERENCE: 10466-14
```

```
;  CURRENT APPLICATION NUMBER: US/09/909,204
;  CURRENT FILING DATE:  2001-07-18
;  PRIOR APPLICATION NUMBER: PCT/US00/04414
;  PRIOR FILING DATE: 2000-02-22
;  PRIOR APPLICATION NUMBER: US 60/143,048
;  PRIOR FILING DATE: 1999-07-07
;  PRIOR APPLICATION NUMBER: US 60/145,698
;  PRIOR FILING DATE: 1999-07-26
;  PRIOR APPLICATION NUMBER: US 60/146,222
;  PRIOR FILING DATE: 1999-07-28
;  PRIOR APPLICATION NUMBER: PCT/US99/20594
;  PRIOR FILING DATE: 1999-09-08
;  PRIOR APPLICATION NUMBER: PCT/US99/20944
;  PRIOR FILING DATE: 1999-09-13
;  PRIOR APPLICATION NUMBER: PCT/US99/21090
;  PRIOR FILING DATE: 1999-09-15
;  PRIOR APPLICATION NUMBER: PCT/US99/21547
;  PRIOR FILING DATE: 1999-09-15
;  PRIOR APPLICATION NUMBER: PCT/US99/23089
;  PRIOR FILING DATE: 1999-10-05
;  PRIOR APPLICATION NUMBER: PCT/US99/28214
;  PRIOR FILING DATE: 1999-11-29
;  PRIOR APPLICATION NUMBER: PCT/US99/28313
;  PRIOR FILING DATE: 1999-11-30
;  PRIOR APPLICATION NUMBER: PCT/US99/28564
;  PRIOR FILING DATE: 1999-12-02
;  PRIOR APPLICATION NUMBER: PCT/US99/28565
;  PRIOR FILING DATE: 1999-12-02
;  PRIOR APPLICATION NUMBER: PCT/US99/30095
;  PRIOR FILING DATE: 1999-12-16
;  PRIOR APPLICATION NUMBER: PCT/US99/30911
;  PRIOR FILING DATE: 1999-12-20
;  PRIOR APPLICATION NUMBER: PCT/US99/30999
;  PRIOR FILING DATE: 1999-12-20
;  PRIOR APPLICATION NUMBER: PCT/US00/00219
;  PRIOR FILING DATE: 2000-01-05
;  NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 137
;   LENGTH: 382
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-909-204-137

  Query Match             98.7%;  Score 1101;  DB 10;  Length 382;
  Best Local Similarity   96.3%;  Pred. No. 8.8e-103;
  Matches  206;  Conservative    0;  Mismatches    0;  Indels    8;  Gaps    1;

Qy          1 ATGRLLS--------GQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 52
              |||||||        ||||||||||||||||||||||||||||||||||||||||||||
Db         22 ATGRLLSASDLDLRGGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 81

Qy         53 ESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVD 112
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 ESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVD 141

Qy        113 EPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGE 172
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 EPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGE 201

Qy        173 ETELTTPVLPEETQEEDAKKTFKESREAALNLAY 206
              |||||||||||||||||||||||||||||||||
Db        202 ETELTTPVLPEETQEEDAKKTFKESREAALNLAY 235


RESULT 23
US-09-904-820-137
; Sequence 137, Application US/09904820
; Publication No. US20030036094A1
; GENERAL INFORMATION:
;  APPLICANT: Genentech, Inc.
;  APPLICANT:  Ashkenazi, Avi
;  APPLICANT:  Botstein, David
;  APPLICANT:  Desnoyers, Luc
;  APPLICANT:  Eaton, Dan L.
;  APPLICANT:  Ferrara, Napoleone
;  APPLICANT:  Filvaroff, Ellen
;  APPLICANT:  Fong, Sherman
;  APPLICANT:  Gao, Wei-Qiang
;  APPLICANT:  Gerber, Hanspeter
;  APPLICANT:  Gerritsen, Mary E.
;  APPLICANT:  Goddard, A.
;  APPLICANT:  Godowski, Paul J.
;  APPLICANT:  Grimaldi, Christopher J.
;  APPLICANT:  Gurney, Austin L.
;  APPLICANT:  Hillan, Kenneth, J.
;  APPLICANT:  Kljavin, Ivar J.
;  APPLICANT:  Mather, Jennie P.
;  APPLICANT:  Pan, James
;  APPLICANT:  Paoni, Nicholas F.
;  APPLICANT:  Roy, Margaret Ann
;  APPLICANT:  Stewart, Timothy A.
;  APPLICANT:  Tumas, Daniel
;  APPLICANT:  Williams, P. Mickey
;  APPLICANT:  Wood, William, I.
;  TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
;  TITLE OF INVENTION:  Acids Encoding the Same
;  FILE REFERENCE: 10466-14
;  CURRENT APPLICATION NUMBER: US/09/904,820
;  CURRENT FILING DATE:  2001-07-13
;  PRIOR APPLICATION NUMBER: 09/665,350
;  PRIOR FILING DATE: 2000-09-18
;  PRIOR APPLICATION NUMBER: PCT/US00/04414
;  PRIOR FILING DATE: 2000-02-22
;  PRIOR APPLICATION NUMBER: US 60/143,048
;  PRIOR FILING DATE: 1999-07-07
;  PRIOR APPLICATION NUMBER: US 60/145,698
;  PRIOR FILING DATE: 1999-07-26
;  PRIOR APPLICATION NUMBER: US 60/146,222
;  PRIOR FILING DATE: 1999-07-28
;  PRIOR APPLICATION NUMBER: PCT/US99/20594
;  PRIOR FILING DATE: 1999-09-08
;  PRIOR APPLICATION NUMBER: PCT/US99/20944
;  PRIOR FILING DATE: 1999-09-13
;  PRIOR APPLICATION NUMBER: PCT/US99/21090
;  PRIOR FILING DATE: 1999-09-15
;  PRIOR APPLICATION NUMBER: PCT/US99/21547
;  PRIOR FILING DATE: 1999-09-15
;  PRIOR APPLICATION NUMBER: PCT/US99/23089
;  PRIOR FILING DATE: 1999-10-05
;  PRIOR APPLICATION NUMBER: PCT/US99/28214
;  PRIOR FILING DATE: 1999-11-29
;  PRIOR APPLICATION NUMBER: PCT/US99/28313
;  PRIOR FILING DATE: 1999-11-30
;  PRIOR APPLICATION NUMBER: PCT/US99/28564
;  PRIOR FILING DATE: 1999-12-02
;  PRIOR APPLICATION NUMBER: PCT/US99/28565
;  PRIOR FILING DATE: 1999-12-02
;  PRIOR APPLICATION NUMBER: PCT/US99/30095
;  PRIOR FILING DATE: 1999-12-16
;  PRIOR APPLICATION NUMBER: PCT/US99/30911
;  PRIOR FILING DATE: 1999-12-20
;  PRIOR APPLICATION NUMBER: PCT/US99/30999
;  PRIOR FILING DATE: 1999-12-20
;  PRIOR APPLICATION NUMBER: PCT/US00/00219
;  PRIOR FILING DATE: 2000-01-05
;  NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 137
;   LENGTH: 382
;   TYPE: PRT
;   ORGANISM: Homo Sapien
US-09-904-820-137

  Query Match             98.7%;  Score 1101;  DB 10;  Length 382;
  Best Local Similarity   96.3%;  Pred. No. 8.8e-103;
  Matches  206;  Conservative    0;  Mismatches    0;  Indels    8;  Gaps    1;
```

```
Qy          1 ATGRLLS--------GQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 52
              |||||||        ||||||||||||||||||||||||||||||||||||||||||||
Db         22 ATGRLLSASDLDLRGGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 81

Qy         53 ESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVD 112
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 ESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVD 141

Qy        113 EPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGE 172
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 EPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGE 201

Qy        173 ETELTTPVLPEETQEEDAKKTFKESREAALNLAY 206
              |||||||||||||||||||||||||||||||||
Db        202 ETELTTPVLPEETQEEDAKKTFKESREAALNLAY 235
```

RESULT 24
US-09-904-786-137
; Sequence 137, Application US/09904786
; Publication No. US20030039969A1
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/904,786
; CURRENT FILING DATE: 2001-07-12
; PRIOR APPLICATION NUMBER: 09/665,350
; PRIOR FILING DATE: 2000-09-18
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 137
; LENGTH: 382
; TYPE: PRT
; ORGANISM: Homo Sapien
US-09-904-786-137

Query Match 98.7%; Score 1101; DB 10; Length 382;
Best Local Similarity 96.3%; Pred. No. 8.8e-103;
Matches 206; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

```
Qy          1 ATGRLLS--------GQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 52
              |||||||        ||||||||||||||||||||||||||||||||||||||||||||
Db         22 ATGRLLSASDLDLRGGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 81

Qy         53 ESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVD 112
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 ESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVD 141

Qy        113 EPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGE 172
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 EPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGE 201

Qy        173 ETELTTPVLPEETQEEDAKKTFKESREAALNLAY 206
              |||||||||||||||||||||||||||||||||
Db        202 ETELTTPVLPEETQEEDAKKTFKESREAALNLAY 235
```

RESULT 25
US-09-906-646-137
; Sequence 137, Application US/09906646
; Publication No. US20030039971A1
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/906,646
; CURRENT FILING DATE: 2002-01-22
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911

```
;  PRIOR FILING DATE: 1999-12-20
;  PRIOR APPLICATION NUMBER: PCT/US99/30999
;  PRIOR FILING DATE: 1999-12-20
;  PRIOR APPLICATION NUMBER: PCT/US00/00219
;  PRIOR FILING DATE: 2000-01-05
;  NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 137
;   LENGTH: 382
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-906-646-137

  Query Match           98.7%;  Score 1101;  DB 10;  Length 382;
  Best Local Similarity  96.3%;  Pred. No. 8.8e-103;
  Matches  206;  Conservative    0;  Mismatches    0;  Indels    8;  Gaps    1;

Qy          1 ATGRLLS--------GQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 52
              |||||||        ||||||||||||||||||||||||||||||||||||||||||||
Db         22 ATGRLLSASDLDLRGGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 81

Qy         53 ESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVD 112
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 ESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVD 141

Qy        113 EPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGE 172
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 EPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGE 201

Qy        173 ETELTTPVLPEETQEEDAKKTFKESREAALNLAY 206
              |||||||||||||||||||||||||||||||||
Db        202 ETELTTPVLPEETQEEDAKKTFKESREAALNLAY 235


RESULT 26
US-09-906-700-137
; Sequence 137, Application US/09906700
; Publication No. US20030039972A1
; GENERAL INFORMATION:
;  APPLICANT: Genentech, Inc.
;  APPLICANT:  Ashkenazi, Avi
;  APPLICANT:  Botstein, David
;  APPLICANT:  Desnoyers, Luc
;  APPLICANT:  Eaton, Dan L.
;  APPLICANT:  Ferrara, Napoleone
;  APPLICANT:  Filvaroff, Ellen
;  APPLICANT:  Fong, Sherman
;  APPLICANT:  Gao, Wei-Qiang
;  APPLICANT:  Gerber, Hanspeter
;  APPLICANT:  Gerritsen, Mary E.
;  APPLICANT:  Goddard, A.
;  APPLICANT:  Godowski, Paul J.
;  APPLICANT:  Grimaldi, Christopher J.
;  APPLICANT:  Gurney, Austin L.
;  APPLICANT:  Hillan, Kenneth, J.
;  APPLICANT:  Kljavin, Ivar J.
;  APPLICANT:  Mather, Jennie P.
;  APPLICANT:  Pan, James
;  APPLICANT:  Paoni, Nicholas F.
;  APPLICANT:  Roy, Margaret Ann
;  APPLICANT:  Stewart, Timothy A.
;  APPLICANT:  Tumas, Daniel
;  APPLICANT:  Williams, P. Mickey
;  APPLICANT:  Wood, William, I.
;  TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
;  TITLE OF INVENTION:  Acids Encoding the Same
;  FILE REFERENCE: 10466-14
;  CURRENT APPLICATION NUMBER: US/09/906,700
;  CURRENT FILING DATE:  2000-09-18
;  PRIOR APPLICATION NUMBER: PCT/US00/04414
;  PRIOR FILING DATE: 2000-02-22
;  PRIOR APPLICATION NUMBER: US 60/143,048
;  PRIOR FILING DATE: 1999-07-07
;  PRIOR APPLICATION NUMBER: US 60/145,698
;  PRIOR FILING DATE: 1999-07-26
;  PRIOR APPLICATION NUMBER: US 60/146,222
;  PRIOR FILING DATE: 1999-07-28
;  PRIOR APPLICATION NUMBER: PCT/US99/20594
;  PRIOR FILING DATE: 1999-09-08
;  PRIOR APPLICATION NUMBER: PCT/US99/20944
;  PRIOR FILING DATE: 1999-09-13
;  PRIOR APPLICATION NUMBER: PCT/US99/21090
;  PRIOR FILING DATE: 1999-09-15
;  PRIOR APPLICATION NUMBER: PCT/US99/21547
;  PRIOR FILING DATE: 1999-09-15
;  PRIOR APPLICATION NUMBER: PCT/US99/23089
;  PRIOR FILING DATE: 1999-10-05
;  PRIOR APPLICATION NUMBER: PCT/US99/28214
;  PRIOR FILING DATE: 1999-11-29
;  PRIOR APPLICATION NUMBER: PCT/US99/28313
;  PRIOR FILING DATE: 1999-11-30
;  PRIOR APPLICATION NUMBER: PCT/US99/28564
;  PRIOR FILING DATE: 1999-12-02
;  PRIOR APPLICATION NUMBER: PCT/US99/28565
;  PRIOR FILING DATE: 1999-12-02
;  PRIOR APPLICATION NUMBER: PCT/US99/30095
;  PRIOR FILING DATE: 1999-12-16
;  PRIOR APPLICATION NUMBER: PCT/US99/30911
;  PRIOR FILING DATE: 1999-12-20
;  PRIOR APPLICATION NUMBER: PCT/US99/30999
;  PRIOR FILING DATE: 1999-12-20
;  PRIOR APPLICATION NUMBER: PCT/US00/00219
;  PRIOR FILING DATE: 2000-01-05
;  NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 137
;   LENGTH: 382
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-906-700-137

  Query Match           98.7%;  Score 1101;  DB 10;  Length 382;
  Best Local Similarity  96.3%;  Pred. No. 8.8e-103;
  Matches  206;  Conservative    0;  Mismatches    0;  Indels    8;  Gaps    1;

Qy          1 ATGRLLS--------GQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 52
              |||||||        ||||||||||||||||||||||||||||||||||||||||||||
Db         22 ATGRLLSASDLDLRGGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 81

Qy         53 ESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVD 112
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 ESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVD 141

Qy        113 EPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGE 172
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 EPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGE 201

Qy        173 ETELTTPVLPEETQEEDAKKTFKESREAALNLAY 206
              |||||||||||||||||||||||||||||||||
Db        202 ETELTTPVLPEETQEEDAKKTFKESREAALNLAY 235


RESULT 27
US-09-903-786-137
; Sequence 137, Application US/09903786
; Publication No. US20030044793A1
; GENERAL INFORMATION:
;  APPLICANT: Genentech, Inc.
;  APPLICANT:  Ashkenazi, Avi
;  APPLICANT:  Botstein, David
;  APPLICANT:  Desnoyers, Luc
;  APPLICANT:  Eaton, Dan L.
;  APPLICANT:  Ferrara, Napoleone
;  APPLICANT:  Filvaroff, Ellen
;  APPLICANT:  Fong, Sherman
;  APPLICANT:  Gao, Wei-Qiang
;  APPLICANT:  Gerber, Hanspeter
;  APPLICANT:  Gerritsen, Mary E.
```

```
;  APPLICANT:  Goddard, A.
;  APPLICANT:  Godowski, Paul J.
;  APPLICANT:  Grimaldi, Christopher J.
;  APPLICANT:  Gurney, Austin L.
;  APPLICANT:  Hillan, Kenneth, J.
;  APPLICANT:  Kljavin, Ivar J.
;  APPLICANT:  Mather, Jennie P.
;  APPLICANT:  Pan, James
;  APPLICANT:  Paoni, Nicholas F.
;  APPLICANT:  Roy, Margaret Ann
;  APPLICANT:  Stewart, Timothy A.
;  APPLICANT:  Tumas, Daniel
;  APPLICANT:  Williams, P. Mickey
;  APPLICANT:  Wood, William, I.
;  TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
;  TITLE OF INVENTION:  Acids Encoding the Same
;  FILE REFERENCE: 10466-14
;  CURRENT APPLICATION NUMBER: US/09/903,786
;  CURRENT FILING DATE:  2001-07-11
;  PRIOR APPLICATION NUMBER: 09/665,350
;  PRIOR FILING DATE: 2000-09-18
;  PRIOR APPLICATION NUMBER: PCT/US00/04414
;  PRIOR FILING DATE: 2000-02-22
;  PRIOR APPLICATION NUMBER: US 60/143,048
;  PRIOR FILING DATE: 1999-07-07
;  PRIOR APPLICATION NUMBER: US 60/145,698
;  PRIOR FILING DATE: 1999-07-26
;  PRIOR APPLICATION NUMBER: US 60/146,222
;  PRIOR FILING DATE: 1999-07-28
;  PRIOR APPLICATION NUMBER: PCT/US99/20594
;  PRIOR FILING DATE: 1999-09-08
;  PRIOR APPLICATION NUMBER: PCT/US99/20944
;  PRIOR FILING DATE: 1999-09-13
;  PRIOR APPLICATION NUMBER: PCT/US99/21090
;  PRIOR FILING DATE: 1999-09-15
;  PRIOR APPLICATION NUMBER: PCT/US99/21547
;  PRIOR FILING DATE: 1999-09-15
;  PRIOR APPLICATION NUMBER: PCT/US99/23089
;  PRIOR FILING DATE: 1999-10-05
;  PRIOR APPLICATION NUMBER: PCT/US99/28214
;  PRIOR FILING DATE: 1999-11-29
;  PRIOR APPLICATION NUMBER: PCT/US99/28313
;  PRIOR FILING DATE: 1999-11-30
;  PRIOR APPLICATION NUMBER: PCT/US99/28564
;  PRIOR FILING DATE: 1999-12-02
;  PRIOR APPLICATION NUMBER: PCT/US99/28565
;  PRIOR FILING DATE: 1999-12-02
;  PRIOR APPLICATION NUMBER: PCT/US99/30095
;  PRIOR FILING DATE: 1999-12-16
;  PRIOR APPLICATION NUMBER: PCT/US99/30911
;  PRIOR FILING DATE: 1999-12-20
;  PRIOR APPLICATION NUMBER: PCT/US99/30999
;  PRIOR FILING DATE: 1999-12-20
;  PRIOR APPLICATION NUMBER: PCT/US00/00219
;  PRIOR FILING DATE: 2000-01-05
;  NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 137
;   LENGTH: 382
;   TYPE: PRT
;   ORGANISM: Homo Sapien
US-09-903-786-137

  Query Match             98.7%;  Score 1101;  DB 10;  Length 382;
  Best Local Similarity   96.3%;  Pred. No. 8.8e-103;
  Matches  206;  Conservative    0;  Mismatches    0;  Indels    8;  Gaps    1;

Qy          1 ATGRLLS--------GQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI  52
              |||||||        ||||||||||||||||||||||||||||||||||||||||||||
Db         22 ATGRLLSASDLDLRGGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI  81

Qy         53 ESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVD  112
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 ESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVD  141

Qy        113 EPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGE  172
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 EPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGE  201

Qy        173 ETELTTPVLPEETQEEDAKKTFKESREAALNLAY  206
              |||||||||||||||||||||||||||||||||
Db        202 ETELTTPVLPEETQEEDAKKTFKESREAALNLAY  235


RESULT 28
US-09-902-903-137
; Sequence 137, Application US/09902903
; Publication No. US20030044839A1
; GENERAL INFORMATION:
;  APPLICANT: Genentech, Inc.
;  APPLICANT:  Ashkenazi, Avi
;  APPLICANT:  Botstein, David
;  APPLICANT:  Desnoyers, Luc
;  APPLICANT:  Eaton, Dan L.
;  APPLICANT:  Ferrara, Napoleone
;  APPLICANT:  Filvaroff, Ellen
;  APPLICANT:  Fong, Sherman
;  APPLICANT:  Gao, Wei-Qiang
;  APPLICANT:  Gerber, Hanspeter
;  APPLICANT:  Gerritsen, Mary E.
;  APPLICANT:  Goddard, A.
;  APPLICANT:  Godowski, Paul J.
;  APPLICANT:  Grimaldi, Christopher J.
;  APPLICANT:  Gurney, Austin L.
;  APPLICANT:  Hillan, Kenneth, J.
;  APPLICANT:  Kljavin, Ivar J.
;  APPLICANT:  Mather, Jennie P.
;  APPLICANT:  Pan, James
;  APPLICANT:  Paoni, Nicholas F.
;  APPLICANT:  Roy, Margaret Ann
;  APPLICANT:  Stewart, Timothy A.
;  APPLICANT:  Tumas, Daniel
;  APPLICANT:  Williams, P. Mickey
;  APPLICANT:  Wood, William, I.
;  TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
;  TITLE OF INVENTION:  Acids Encoding the Same
;  FILE REFERENCE: 10466-14
;  CURRENT APPLICATION NUMBER: US/09/902,903
;  CURRENT FILING DATE:  2001-07-10
;  PRIOR APPLICATION NUMBER: PCT/US00/04414
;  PRIOR FILING DATE: 2000-02-22
;  PRIOR APPLICATION NUMBER: US 60/143,048
;  PRIOR FILING DATE: 1999-07-07
;  PRIOR APPLICATION NUMBER: US 60/145,698
;  PRIOR FILING DATE: 1999-07-26
;  PRIOR APPLICATION NUMBER: US 60/146,222
;  PRIOR FILING DATE: 1999-07-28
;  PRIOR APPLICATION NUMBER: PCT/US99/20594
;  PRIOR FILING DATE: 1999-09-08
;  PRIOR APPLICATION NUMBER: PCT/US99/20944
;  PRIOR FILING DATE: 1999-09-13
;  PRIOR APPLICATION NUMBER: PCT/US99/21090
;  PRIOR FILING DATE: 1999-09-15
;  PRIOR APPLICATION NUMBER: PCT/US99/21547
;  PRIOR FILING DATE: 1999-09-15
;  PRIOR APPLICATION NUMBER: PCT/US99/23089
;  PRIOR FILING DATE: 1999-10-05
;  PRIOR APPLICATION NUMBER: PCT/US99/28214
;  PRIOR FILING DATE: 1999-11-29
;  PRIOR APPLICATION NUMBER: PCT/US99/28313
;  PRIOR FILING DATE: 1999-11-30
;  PRIOR APPLICATION NUMBER: PCT/US99/28564
;  PRIOR FILING DATE: 1999-12-02
;  PRIOR APPLICATION NUMBER: PCT/US99/28565
;  PRIOR FILING DATE: 1999-12-02
;  PRIOR APPLICATION NUMBER: PCT/US99/30095
;  PRIOR FILING DATE: 1999-12-16
```

```
;  PRIOR APPLICATION NUMBER: PCT/US99/30911
;  PRIOR FILING DATE: 1999-12-20
;  PRIOR APPLICATION NUMBER: PCT/US99/30999
;  PRIOR FILING DATE: 1999-12-20
;  PRIOR APPLICATION NUMBER: PCT/US00/00219
;  PRIOR FILING DATE: 2000-01-05
;  NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 137
;   LENGTH: 382
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-902-903-137

  Query Match             98.7%;  Score 1101;  DB 10;  Length 382;
  Best Local Similarity   96.3%;  Pred. No. 8.8e-103;
  Matches  206;  Conservative    0;  Mismatches    0;  Indels    8;  Gaps    1;

Qy          1 ATGRLLS--------GQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 52
              |||||||        ||||||||||||||||||||||||||||||||||||||||||||
Db         22 ATGRLLSASDLDLRGGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 81

Qy         53 ESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVD 112
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 ESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVD 141

Qy        113 EPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGE 172
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 EPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGE 201

Qy        173 ETELTTPVLPEETQEEDAKKTFKESREAALNLAY 206
              |||||||||||||||||||||||||||||||||
Db        202 ETELTTPVLPEETQEEDAKKTFKESREAALNLAY 235


RESULT 29
US-09-903-749A-137
; Sequence 137, Application US/09903749A
; Publication No. US20030045693A1
; GENERAL INFORMATION:
;  APPLICANT: Genentech, Inc.
;  APPLICANT:  Ashkenazi, Avi
;  APPLICANT:  Botstein, David
;  APPLICANT:  Desnoyers, Luc
;  APPLICANT:  Eaton, Dan L.
;  APPLICANT:  Ferrara, Napoleone
;  APPLICANT:  Filvaroff, Ellen
;  APPLICANT:  Fong, Sherman
;  APPLICANT:  Gao, Wei-Qiang
;  APPLICANT:  Gerber, Hanspeter
;  APPLICANT:  Gerritsen, Mary E.
;  APPLICANT:  Goddard, A.
;  APPLICANT:  Godowski, Paul J.
;  APPLICANT:  Grimaldi, Christopher J.
;  APPLICANT:  Gurney, Austin L.
;  APPLICANT:  Hillan, Kenneth, J.
;  APPLICANT:  Kljavin, Ivar J.
;  APPLICANT:  Mather, Jennie P.
;  APPLICANT:  Pan, James
;  APPLICANT:  Paoni, Nicholas F.
;  APPLICANT:  Roy, Margaret Ann
;  APPLICANT:  Stewart, Timothy A.
;  APPLICANT:  Tumas, Daniel
;  APPLICANT:  Williams, P. Mickey
;  APPLICANT:  Wood, William, I.
;  TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
;  TITLE OF INVENTION:  Acids Encoding the Same
;  FILE REFERENCE: 10466-14
;  CURRENT APPLICATION NUMBER: US/09/903,749A
;  CURRENT FILING DATE:  2001-07-11
;  PRIOR APPLICATION NUMBER: PCT/US00/04414
;  PRIOR FILING DATE: 2000-02-22
;  PRIOR APPLICATION NUMBER: US 60/143,048
;  PRIOR FILING DATE: 1999-07-07
;  PRIOR APPLICATION NUMBER: US 60/145,698
;  PRIOR FILING DATE: 1999-07-26
;  PRIOR APPLICATION NUMBER: US 60/146,222
;  PRIOR FILING DATE: 1999-07-28
;  PRIOR APPLICATION NUMBER: PCT/US99/20594
;  PRIOR FILING DATE: 1999-09-08
;  PRIOR APPLICATION NUMBER: PCT/US99/20944
;  PRIOR FILING DATE: 1999-09-13
;  PRIOR APPLICATION NUMBER: PCT/US99/21090
;  PRIOR FILING DATE: 1999-09-15
;  PRIOR APPLICATION NUMBER: PCT/US99/21547
;  PRIOR FILING DATE: 1999-09-15
;  PRIOR APPLICATION NUMBER: PCT/US99/23089
;  PRIOR FILING DATE: 1999-10-05
;  PRIOR APPLICATION NUMBER: PCT/US99/28214
;  PRIOR FILING DATE: 1999-11-29
;  PRIOR APPLICATION NUMBER: PCT/US99/28313
;  PRIOR FILING DATE: 1999-11-30
;  PRIOR APPLICATION NUMBER: PCT/US99/28564
;  PRIOR FILING DATE: 1999-12-02
;  PRIOR APPLICATION NUMBER: PCT/US99/28565
;  PRIOR FILING DATE: 1999-12-02
;  PRIOR APPLICATION NUMBER: PCT/US99/30095
;  PRIOR FILING DATE: 1999-12-16
;  PRIOR APPLICATION NUMBER: PCT/US99/30911
;  PRIOR FILING DATE: 1999-12-20
;  PRIOR APPLICATION NUMBER: PCT/US99/30999
;  PRIOR FILING DATE: 1999-12-20
;  PRIOR APPLICATION NUMBER: PCT/US00/00219
;  PRIOR FILING DATE: 2000-01-05
;  NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 137
;   LENGTH: 382
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-903-749A-137

  Query Match             98.7%;  Score 1101;  DB 10;  Length 382;
  Best Local Similarity   96.3%;  Pred. No. 8.8e-103;
  Matches  206;  Conservative    0;  Mismatches    0;  Indels    8;  Gaps    1;

Qy          1 ATGRLLS--------GQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 52
              |||||||        ||||||||||||||||||||||||||||||||||||||||||||
Db         22 ATGRLLSASDLDLRGGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 81

Qy         53 ESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVD 112
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 ESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVD 141

Qy        113 EPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGE 172
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 EPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGE 201

Qy        173 ETELTTPVLPEETQEEDAKKTFKESREAALNLAY 206
              |||||||||||||||||||||||||||||||||
Db        202 ETELTTPVLPEETQEEDAKKTFKESREAALNLAY 235


RESULT 30
US-09-904-119-137
; Sequence 137, Application US/09904119
; Publication No. US20030049621A1
; GENERAL INFORMATION:
;  APPLICANT: Genentech, Inc.
;  APPLICANT:  Ashkenazi, Avi
;  APPLICANT:  Botstein, David
;  APPLICANT:  Desnoyers, Luc
;  APPLICANT:  Eaton, Dan L.
;  APPLICANT:  Ferrara, Napoleone
;  APPLICANT:  Filvaroff, Ellen
;  APPLICANT:  Fong, Sherman
;  APPLICANT:  Gao, Wei-Qiang
;  APPLICANT:  Gerber, Hanspeter
```

; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/904,119
; CURRENT FILING DATE: 2001-07-11
; PRIOR APPLICATION NUMBER: 09/665,350
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 137
; LENGTH: 382
; TYPE: PRT
; ORGANISM: Homo Sapien
US-09-904-119-137

Query Match 98.7%; Score 1101; DB 10; Length 382;
Best Local Similarity 96.3%; Pred. No. 8.8e-103;
Matches 206; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

```
Qy          1 ATGRLLS--------GQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI  52
              |||||||        ||||||||||||||||||||||||||||||||||||||||||||
Db         22 ATGRLLSASDLDLRGGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI  81

Qy         53 ESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVD  112
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 ESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVD  141

Qy        113 EPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGE  172
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 EPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGE  201

Qy        173 ETELTTPVLPEETQEEDAKKTFKESREAALNLAY  206
              |||||||||||||||||||||||||||||||||
Db        202 ETELTTPVLPEETQEEDAKKTFKESREAALNLAY  235
```

RESULT 31
US-09-904-956-137
; Sequence 137, Application US/09904956
; Publication No. US20030049622A1
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/904,956
; CURRENT FILING DATE: 2001-07-12
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095

; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 137
; LENGTH: 382
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-904-956-137

Query Match 98.7%; Score 1101; DB 10; Length 382;
Best Local Similarity 96.3%; Pred. No. 8.8e-103;
Matches 206; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

```
Qy          1 ATGRLLS--------GQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 52
              |||||||        ||||||||||||||||||||||||||||||||||||||||||||
Db         22 ATGRLLSASDLDLRGGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 81

Qy         53 ESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVD 112
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 ESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVD 141

Qy        113 EPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGE 172
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 EPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGE 201

Qy        173 ETELTTPVLPEETQEEDAKKTFKESREAALNLAY 206
              |||||||||||||||||||||||||||||||||
Db        202 ETELTTPVLPEETQEEDAKKTFKESREAALNLAY 235
```

RESULT 32
US-09-902-736-137
; Sequence 137, Application US/09902736
; Publication No. US20030049676A1
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/902,736
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 09/665,350
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 137
; LENGTH: 382
; TYPE: PRT
; ORGANISM: Homo Sapien
US-09-902-736-137

Query Match 98.7%; Score 1101; DB 10; Length 382;
Best Local Similarity 96.3%; Pred. No. 8.8e-103;
Matches 206; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

```
Qy          1 ATGRLLS--------GQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 52
              |||||||        ||||||||||||||||||||||||||||||||||||||||||||
Db         22 ATGRLLSASDLDLRGGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 81

Qy         53 ESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVD 112
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 ESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVD 141

Qy        113 EPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGE 172
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 EPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGE 201

Qy        173 ETELTTPVLPEETQEEDAKKTFKESREAALNLAY 206
              |||||||||||||||||||||||||||||||||
Db        202 ETELTTPVLPEETQEEDAKKTFKESREAALNLAY 235
```

RESULT 33
US-09-907-794-137
; Sequence 137, Application US/09907794
; Publication No. US20030049677A1
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen

; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/907,794
; CURRENT FILING DATE: 2001-07-17
; PRIOR APPLICATION NUMBER: 09/665,350
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 137
; LENGTH: 382
; TYPE: PRT
; ORGANISM: Homo Sapien
US-09-907-794-137

Query Match 98.7%; Score 1101; DB 10; Length 382;
Best Local Similarity 96.3%; Pred. No. 8.8e-103;
Matches 206; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

```
Qy          1 ATGRLLS--------GQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 52
              |||||||        ||||||||||||||||||||||||||||||||||||||||||||
Db         22 ATGRLLSASDLDLRGGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 81

Qy         53 ESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVD 112
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 ESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVD 141

Qy        113 EPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGE 172
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 EPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGE 201

Qy        173 ETELTTPVLPEETQEEDAKKTFKESREAALNLAY 206
              |||||||||||||||||||||||||||||||||
Db        202 ETELTTPVLPEETQEEDAKKTFKESREAALNLAY 235
```

RESULT 34
US-09-903-943-137
; Sequence 137, Application US/09903943
; Publication No. US20030054349A1
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/903,943
; CURRENT FILING DATE: 2001-07-11
; PRIOR APPLICATION NUMBER: 09/665,350
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30

; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 137
; LENGTH: 382
; TYPE: PRT
; ORGANISM: Homo Sapien
US-09-903-943-137

Query Match 98.7%; Score 1101; DB 10; Length 382;
Best Local Similarity 96.3%; Pred. No. 8.8e-103;
Matches 206; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

```
Qy          1 ATGRLLS--------GQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 52
              |||||||        ||||||||||||||||||||||||||||||||||||||||||||
Db         22 ATGRLLSASDLDLRGGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 81

Qy         53 ESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVD 112
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 ESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVD 141

Qy        113 EPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGE 172
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 EPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGE 201

Qy        173 ETELTTPVLPEETQEEDAKKTFKESREAALNLAY 206
              |||||||||||||||||||||||||||||||||
Db        202 ETELTTPVLPEETQEEDAKKTFKESREAALNLAY 235
```

RESULT 35
US-09-904-462-137
; Sequence 137, Application US/09904462
; Publication No. US20030054351A1
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/904,462
; CURRENT FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: 09/665,350
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 137
; LENGTH: 382
; TYPE: PRT
; ORGANISM: Homo Sapien
US-09-904-462-137

Query Match 98.7%; Score 1101; DB 10; Length 382;
Best Local Similarity 96.3%; Pred. No. 8.8e-103;
Matches 206; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

```
Qy          1 ATGRLLS--------GQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 52
              |||||||        ||||||||||||||||||||||||||||||||||||||||||||
Db         22 ATGRLLSASDLDLRGGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 81

Qy         53 ESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVD 112
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 ESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVD 141

Qy        113 EPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGE 172
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 EPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGE 201

Qy        173 ETELTTPVLPEETQEEDAKKTFKESREAALNLAY 206
              |||||||||||||||||||||||||||||||||
Db        202 ETELTTPVLPEETQEEDAKKTFKESREAALNLAY 235
```

RESULT 36
US-09-907-925-137
; Sequence 137, Application US/09907925
; Publication No. US20030054352A1
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi

```
;  APPLICANT:  Botstein, David
;  APPLICANT:  Desnoyers, Luc
;  APPLICANT:  Eaton, Dan L.
;  APPLICANT:  Ferrara, Napoleone
;  APPLICANT:  Filvaroff, Ellen
;  APPLICANT:  Fong, Sherman
;  APPLICANT:  Gao, Wei-Qiang
;  APPLICANT:  Gerber, Hanspeter
;  APPLICANT:  Gerritsen, Mary E.
;  APPLICANT:  Goddard, A.
;  APPLICANT:  Godowski, Paul J.
;  APPLICANT:  Grimaldi, Christopher J.
;  APPLICANT:  Gurney, Austin L.
;  APPLICANT:  Hillan, Kenneth, J.
;  APPLICANT:  Kljavin, Ivar J.
;  APPLICANT:  Mather, Jennie P.
;  APPLICANT:  Pan, James
;  APPLICANT:  Paoni, Nicholas F.
;  APPLICANT:  Roy, Margaret Ann
;  APPLICANT:  Stewart, Timothy A.
;  APPLICANT:  Tumas, Daniel
;  APPLICANT:  Williams, P. Mickey
;  APPLICANT:  Wood, William, I.
;  TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
;  TITLE OF INVENTION:  Acids Encoding the Same
;  FILE REFERENCE: 10466-14
;  CURRENT APPLICATION NUMBER: US/09/907,925
;  CURRENT FILING DATE:  2001-07-17
;  PRIOR APPLICATION NUMBER: 09/665,350
;  PRIOR FILING DATE: 2000-09-18
;  PRIOR APPLICATION NUMBER: PCT/US00/04414
;  PRIOR FILING DATE: 2000-02-22
;  PRIOR APPLICATION NUMBER: US 60/143,048
;  PRIOR FILING DATE: 1999-07-07
;  PRIOR APPLICATION NUMBER: US 60/145,698
;  PRIOR FILING DATE: 1999-07-26
;  PRIOR APPLICATION NUMBER: US 60/146,222
;  PRIOR FILING DATE: 1999-07-28
;  PRIOR APPLICATION NUMBER: PCT/US99/20594
;  PRIOR FILING DATE: 1999-09-08
;  PRIOR APPLICATION NUMBER: PCT/US99/20944
;  PRIOR FILING DATE: 1999-09-13
;  PRIOR APPLICATION NUMBER: PCT/US99/21090
;  PRIOR FILING DATE: 1999-09-15
;  PRIOR APPLICATION NUMBER: PCT/US99/21547
;  PRIOR FILING DATE: 1999-09-15
;  PRIOR APPLICATION NUMBER: PCT/US99/23089
;  PRIOR FILING DATE: 1999-10-05
;  PRIOR APPLICATION NUMBER: PCT/US99/28214
;  PRIOR FILING DATE: 1999-11-29
;  PRIOR APPLICATION NUMBER: PCT/US99/28313
;  PRIOR FILING DATE: 1999-11-30
;  PRIOR APPLICATION NUMBER: PCT/US99/28564
;  PRIOR FILING DATE: 1999-12-02
;  PRIOR APPLICATION NUMBER: PCT/US99/28565
;  PRIOR FILING DATE: 1999-12-02
;  PRIOR APPLICATION NUMBER: PCT/US99/30095
;  PRIOR FILING DATE: 1999-12-16
;  PRIOR APPLICATION NUMBER: PCT/US99/30911
;  PRIOR FILING DATE: 1999-12-20
;  PRIOR APPLICATION NUMBER: PCT/US99/30999
;  PRIOR FILING DATE: 1999-12-20
;  PRIOR APPLICATION NUMBER: PCT/US00/00219
;  PRIOR FILING DATE: 2000-01-05
;  NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 137
;    LENGTH: 382
;    TYPE: PRT
;    ORGANISM: Homo Sapien
US-09-907-925-137

  Query Match             98.7%;  Score 1101;  DB 10;  Length 382;
  Best Local Similarity   96.3%;  Pred. No. 8.8e-103;
  Matches  206;  Conservative    0;  Mismatches    0;  Indels    8;  Gaps    1;

Qy          1 ATGRLLS--------GQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 52
              |||||||        ||||||||||||||||||||||||||||||||||||||||||||
Db         22 ATGRLLSASDLDLRGGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 81

Qy         53 ESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVD 112
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 ESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVD 141

Qy        113 EPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGE 172
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 EPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGE 201

Qy        173 ETELTTPVLPEETQEEDAKKTFKESREAALNLAY 206
              |||||||||||||||||||||||||||||||||
Db        202 ETELTTPVLPEETQEEDAKKTFKESREAALNLAY 235


RESULT 37
US-09-902-692-137
; Sequence 137, Application US/09902692
; Publication No. US20030054400A1
; GENERAL INFORMATION:
;  APPLICANT: Genentech, Inc.
;  APPLICANT:  Ashkenazi, Avi
;  APPLICANT:  Botstein, David
;  APPLICANT:  Desnoyers, Luc
;  APPLICANT:  Eaton, Dan L.
;  APPLICANT:  Ferrara, Napoleone
;  APPLICANT:  Filvaroff, Ellen
;  APPLICANT:  Fong, Sherman
;  APPLICANT:  Gao, Wei-Qiang
;  APPLICANT:  Gerber, Hanspeter
;  APPLICANT:  Gerritsen, Mary E.
;  APPLICANT:  Goddard, A.
;  APPLICANT:  Godowski, Paul J.
;  APPLICANT:  Grimaldi, Christopher J.
;  APPLICANT:  Gurney, Austin L.
;  APPLICANT:  Hillan, Kenneth, J.
;  APPLICANT:  Kljavin, Ivar J.
;  APPLICANT:  Mather, Jennie P.
;  APPLICANT:  Pan, James
;  APPLICANT:  Paoni, Nicholas F.
;  APPLICANT:  Roy, Margaret Ann
;  APPLICANT:  Stewart, Timothy A.
;  APPLICANT:  Tumas, Daniel
;  APPLICANT:  Williams, P. Mickey
;  APPLICANT:  Wood, William, I.
;  TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
;  TITLE OF INVENTION:  Acids Encoding the Same
;  FILE REFERENCE: 10466-14
;  CURRENT APPLICATION NUMBER: US/09/902,692
;  CURRENT FILING DATE:  2001-07-10
;  PRIOR APPLICATION NUMBER: PCT/US00/04414
;  PRIOR FILING DATE: 2000-02-22
;  PRIOR APPLICATION NUMBER: US 60/143,048
;  PRIOR FILING DATE: 1999-07-07
;  PRIOR APPLICATION NUMBER: US 60/145,698
;  PRIOR FILING DATE: 1999-07-26
;  PRIOR APPLICATION NUMBER: US 60/146,222
;  PRIOR FILING DATE: 1999-07-28
;  PRIOR APPLICATION NUMBER: PCT/US99/20594
;  PRIOR FILING DATE: 1999-09-08
;  PRIOR APPLICATION NUMBER: PCT/US99/20944
;  PRIOR FILING DATE: 1999-09-13
;  PRIOR APPLICATION NUMBER: PCT/US99/21090
;  PRIOR FILING DATE: 1999-09-15
;  PRIOR APPLICATION NUMBER: PCT/US99/21547
;  PRIOR FILING DATE: 1999-09-15
;  PRIOR APPLICATION NUMBER: PCT/US99/23089
;  PRIOR FILING DATE: 1999-10-05
;  PRIOR APPLICATION NUMBER: PCT/US99/28214
```

```
;  PRIOR FILING DATE: 1999-11-29
;  PRIOR APPLICATION NUMBER: PCT/US99/28313
;  PRIOR FILING DATE: 1999-11-30
;  PRIOR APPLICATION NUMBER: PCT/US99/28564
;  PRIOR FILING DATE: 1999-12-02
;  PRIOR APPLICATION NUMBER: PCT/US99/28565
;  PRIOR FILING DATE: 1999-12-02
;  PRIOR APPLICATION NUMBER: PCT/US99/30095
;  PRIOR FILING DATE: 1999-12-16
;  PRIOR APPLICATION NUMBER: PCT/US99/30911
;  PRIOR FILING DATE: 1999-12-20
;  PRIOR APPLICATION NUMBER: PCT/US99/30999
;  PRIOR FILING DATE: 1999-12-20
;  PRIOR APPLICATION NUMBER: PCT/US00/00219
;  PRIOR FILING DATE: 2000-01-05
;  NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 137
;   LENGTH: 382
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-902-692-137

  Query Match             98.7%;  Score 1101;  DB 10;  Length 382;
  Best Local Similarity   96.3%;  Pred. No. 8.8e-103;
  Matches  206;  Conservative    0;  Mismatches    0;  Indels    8;  Gaps    1;

Qy          1 ATGRLLS--------GQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 52
              |||||||        ||||||||||||||||||||||||||||||||||||||||||||
Db         22 ATGRLLSASDLDLRGGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 81

Qy         53 ESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVD 112
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 ESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVD 141

Qy        113 EPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGE 172
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 EPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGE 201

Qy        173 ETELTTPVLPEETQEEDAKKTFKESREAALNLAY 206
              |||||||||||||||||||||||||||||||||
Db        202 ETELTTPVLPEETQEEDAKKTFKESREAALNLAY 235


RESULT 38
US-09-903-520-137
; Sequence 137, Application US/09903520
; Publication No. US20030054401A1
; GENERAL INFORMATION:
;  APPLICANT: Genentech, Inc.
;  APPLICANT:  Ashkenazi, Avi
;  APPLICANT:  Botstein, David
;  APPLICANT:  Desnoyers, Luc
;  APPLICANT:  Eaton, Dan L.
;  APPLICANT:  Ferrara, Napoleone
;  APPLICANT:  Filvaroff, Ellen
;  APPLICANT:  Fong, Sherman
;  APPLICANT:  Gao, Wei-Qiang
;  APPLICANT:  Gerber, Hanspeter
;  APPLICANT:  Gerritsen, Mary E.
;  APPLICANT:  Goddard, A.
;  APPLICANT:  Godowski, Paul J.
;  APPLICANT:  Grimaldi, Christopher J.
;  APPLICANT:  Gurney, Austin L.
;  APPLICANT:  Hillan, Kenneth, J.
;  APPLICANT:  Kljavin, Ivar J.
;  APPLICANT:  Mather, Jennie P.
;  APPLICANT:  Pan, James
;  APPLICANT:  Paoni, Nicholas F.
;  APPLICANT:  Roy, Margaret Ann
;  APPLICANT:  Stewart, Timothy A.
;  APPLICANT:  Tumas, Daniel
;  APPLICANT:  Williams, P. Mickey
;  APPLICANT:  Wood, William, I.
;  TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
;  TITLE OF INVENTION:  Acids Encoding the Same
;  FILE REFERENCE: 10466-14
;  CURRENT APPLICATION NUMBER: US/09/903,520
;  CURRENT FILING DATE:  2001-07-11
;  PRIOR APPLICATION NUMBER: 09/665,350
;  PRIOR FILING DATE: 2000-09-18
;  PRIOR APPLICATION NUMBER: PCT/US00/04414
;  PRIOR FILING DATE: 2000-02-22
;  PRIOR APPLICATION NUMBER: US 60/143,048
;  PRIOR FILING DATE: 1999-07-07
;  PRIOR APPLICATION NUMBER: US 60/145,698
;  PRIOR FILING DATE: 1999-07-26
;  PRIOR APPLICATION NUMBER: US 60/146,222
;  PRIOR FILING DATE: 1999-07-28
;  PRIOR APPLICATION NUMBER: PCT/US99/20594
;  PRIOR FILING DATE: 1999-09-08
;  PRIOR APPLICATION NUMBER: PCT/US99/20944
;  PRIOR FILING DATE: 1999-09-13
;  PRIOR APPLICATION NUMBER: PCT/US99/21090
;  PRIOR FILING DATE: 1999-09-15
;  PRIOR APPLICATION NUMBER: PCT/US99/21547
;  PRIOR FILING DATE: 1999-09-15
;  PRIOR APPLICATION NUMBER: PCT/US99/23089
;  PRIOR FILING DATE: 1999-10-05
;  PRIOR APPLICATION NUMBER: PCT/US99/28214
;  PRIOR FILING DATE: 1999-11-29
;  PRIOR APPLICATION NUMBER: PCT/US99/28313
;  PRIOR FILING DATE: 1999-11-30
;  PRIOR APPLICATION NUMBER: PCT/US99/28564
;  PRIOR FILING DATE: 1999-12-02
;  PRIOR APPLICATION NUMBER: PCT/US99/28565
;  PRIOR FILING DATE: 1999-12-02
;  PRIOR APPLICATION NUMBER: PCT/US99/30095
;  PRIOR FILING DATE: 1999-12-16
;  PRIOR APPLICATION NUMBER: PCT/US99/30911
;  PRIOR FILING DATE: 1999-12-20
;  PRIOR APPLICATION NUMBER: PCT/US99/30999
;  PRIOR FILING DATE: 1999-12-20
;  PRIOR APPLICATION NUMBER: PCT/US00/00219
;  PRIOR FILING DATE: 2000-01-05
;  NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 137
;   LENGTH: 382
;   TYPE: PRT
;   ORGANISM: Homo Sapien
US-09-903-520-137

  Query Match             98.7%;  Score 1101;  DB 10;  Length 382;
  Best Local Similarity   96.3%;  Pred. No. 8.8e-103;
  Matches  206;  Conservative    0;  Mismatches    0;  Indels    8;  Gaps    1;

Qy          1 ATGRLLS--------GQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 52
              |||||||        ||||||||||||||||||||||||||||||||||||||||||||
Db         22 ATGRLLSASDLDLRGGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 81

Qy         53 ESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVD 112
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 ESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVD 141

Qy        113 EPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGE 172
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 EPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGE 201

Qy        173 ETELTTPVLPEETQEEDAKKTFKESREAALNLAY 206
              |||||||||||||||||||||||||||||||||
Db        202 ETELTTPVLPEETQEEDAKKTFKESREAALNLAY 235


RESULT 39
US-09-905-056-137
; Sequence 137, Application US/09905056
; Publication No. US20030054441A1
```

; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/905,056
; CURRENT FILING DATE: 2002-01-22
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 137
; LENGTH: 382
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-905-056-137

Query Match 98.7%; Score 1101; DB 10; Length 382;
Best Local Similarity 96.3%; Pred. No. 8.8e-103;
Matches 206; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

```
Qy          1 ATGRLLS--------GQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 52
              |||||||        ||||||||||||||||||||||||||||||||||||||||||||
Db         22 ATGRLLSASDLDLRGGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 81

Qy         53 ESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVD 112
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 ESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVD 141

Qy        113 EPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGE 172
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 EPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGE 201

Qy        173 ETELTTPVLPEETQEEDAKKTFKESREAALNLAY 206
              |||||||||||||||||||||||||||||||||
Db        202 ETELTTPVLPEETQEEDAKKTFKESREAALNLAY 235
```

RESULT 40
US-09-909-064-137
; Sequence 137, Application US/09909064
; Publication No. US20030059772A1
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/909,064
; CURRENT FILING DATE: 2001-07-18
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05

; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 137
; LENGTH: 382
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-909-064-137

Query Match 98.7%; Score 1101; DB 10; Length 382;
Best Local Similarity 96.3%; Pred. No. 8.8e-103;
Matches 206; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

```
Qy          1 ATGRLLS--------GQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 52
              |||||||        ||||||||||||||||||||||||||||||||||||||||||||
Db         22 ATGRLLSASDLDLRGGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 81

Qy         53 ESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVD 112
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 ESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVD 141

Qy        113 EPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGE 172
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 EPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGE 201

Qy        173 ETELTTPVLPEETQEEDAKKTFKESREAALNLAY 206
              |||||||||||||||||||||||||||||||||
Db        202 ETELTTPVLPEETQEEDAKKTFKESREAALNLAY 235
```

RESULT 41
US-09-904-553-137
; Sequence 137, Application US/09904553
; Publication No. US20030059828A1
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/904,553
; CURRENT FILING DATE: 2002-01-22
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 137
; LENGTH: 382
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-904-553-137

Query Match 98.7%; Score 1101; DB 10; Length 382;
Best Local Similarity 96.3%; Pred. No. 8.8e-103;
Matches 206; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

```
Qy          1 ATGRLLS--------GQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 52
              |||||||        ||||||||||||||||||||||||||||||||||||||||||||
Db         22 ATGRLLSASDLDLRGGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 81

Qy         53 ESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVD 112
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 ESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVD 141

Qy        113 EPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGE 172
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 EPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGE 201

Qy        173 ETELTTPVLPEETQEEDAKKTFKESREAALNLAY 206
              |||||||||||||||||||||||||||||||||
Db        202 ETELTTPVLPEETQEEDAKKTFKESREAALNLAY 235
```

RESULT 42
US-09-905-381-137
; Sequence 137, Application US/09905381
; Publication No. US20030059829A1
; GENERAL INFORMATION:

; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/905,381
; CURRENT FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: 09/665,350
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 137
; LENGTH: 382
; TYPE: PRT
; ORGANISM: Homo Sapien
US-09-905-381-137

Query Match 98.7%; Score 1101; DB 10; Length 382;
Best Local Similarity 96.3%; Pred. No. 8.8e-103;
Matches 206; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

```
Qy          1 ATGRLLS--------GQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 52
              |||||||        ||||||||||||||||||||||||||||||||||||||||||||
Db         22 ATGRLLSASDLDLRGGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 81

Qy         53 ESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVD 112
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 ESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVD 141

Qy        113 EPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGE 172
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 EPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGE 201

Qy        173 ETELTTPVLPEETQEEDAKKTFKESREAALNLAY 206
              |||||||||||||||||||||||||||||||||
Db        202 ETELTTPVLPEETQEEDAKKTFKESREAALNLAY 235
```

RESULT 43
US-09-905-088-137
; Sequence 137, Application US/09905088
; Publication No. US20030073077A1
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/905,088
; CURRENT FILING DATE: 2001-07-12
; PRIOR APPLICATION NUMBER: 09/665,350
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547

; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 137
; LENGTH: 382
; TYPE: PRT
; ORGANISM: Homo Sapien
US-09-905-088-137

Query Match 98.7%; Score 1101; DB 10; Length 382;
Best Local Similarity 96.3%; Pred. No. 8.8e-103;
Matches 206; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

```
Qy          1 ATGRLLS--------GQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 52
              |||||||        ||||||||||||||||||||||||||||||||||||||||||||
Db         22 ATGRLLSASDLDLRGGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 81

Qy         53 ESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVD 112
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 ESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVD 141

Qy        113 EPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGE 172
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 EPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGE 201

Qy        173 ETELTTPVLPEETQEEDAKKTFKESREAALNLAY 206
              |||||||||||||||||||||||||||||||||
Db        202 ETELTTPVLPEETQEEDAKKTFKESREAALNLAY 235
```

RESULT 44
US-09-907-575-137
; Sequence 137, Application US/09907575
; Publication No. US20030073079A1
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/907,575
; CURRENT FILING DATE: 2001-12-18
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 137
; LENGTH: 382
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-907-575-137

Query Match 98.7%; Score 1101; DB 10; Length 382;
Best Local Similarity 96.3%; Pred. No. 8.8e-103;
Matches 206; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

```
Qy          1 ATGRLLS--------GQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 52
              |||||||        ||||||||||||||||||||||||||||||||||||||||||||
Db         22 ATGRLLSASDLDLRGGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 81

Qy         53 ESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVD 112
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 ESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVD 141

Qy        113 EPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGE 172
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 EPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGE 201

Qy        173 ETELTTPVLPEETQEEDAKKTFKESREAALNLAY 206
              |||||||||||||||||||||||||||||||||
Db        202 ETELTTPVLPEETQEEDAKKTFKESREAALNLAY 235
```

RESULT 45
US-09-905-075-137

```
; Sequence 137, Application US/09905075
; Publication No. US20030077583A1
; GENERAL INFORMATION:
;  APPLICANT: Genentech, Inc.
;  APPLICANT:  Ashkenazi, Avi
;  APPLICANT:  Botstein, David
;  APPLICANT:  Desnoyers, Luc
;  APPLICANT:  Eaton, Dan L.
;  APPLICANT:  Ferrara, Napoleone
;  APPLICANT:  Filvaroff, Ellen
;  APPLICANT:  Fong, Sherman
;  APPLICANT:  Gao, Wei-Qiang
;  APPLICANT:  Gerber, Hanspeter
;  APPLICANT:  Gerritsen, Mary E.
;  APPLICANT:  Goddard, A.
;  APPLICANT:  Godowski, Paul J.
;  APPLICANT:  Grimaldi, Christopher J.
;  APPLICANT:  Gurney, Austin L.
;  APPLICANT:  Hillan, Kenneth, J.
;  APPLICANT:  Kljavin, Ivar J.
;  APPLICANT:  Mather, Jennie P.
;  APPLICANT:  Pan, James
;  APPLICANT:  Paoni, Nicholas F.
;  APPLICANT:  Roy, Margaret Ann
;  APPLICANT:  Stewart, Timothy A.
;  APPLICANT:  Tumas, Daniel
;  APPLICANT:  Williams, P. Mickey
;  APPLICANT:  Wood, William, I.
;  TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
;  TITLE OF INVENTION:  Acids Encoding the Same
;  FILE REFERENCE: 10466-14
;  CURRENT APPLICATION NUMBER: US/09/905,075
;  CURRENT FILING DATE:  2001-07-13
;  Prior application data removed.  Check file wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 137
;   LENGTH: 382
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-905-075-137

  Query Match             98.7%;  Score 1101;  DB 10;  Length 382;
  Best Local Similarity   96.3%;  Pred. No. 8.8e-103;
  Matches  206;  Conservative    0;  Mismatches    0;  Indels    8;  Gaps    1;

Qy          1 ATGRLLS--------GQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 52
              |||||||        ||||||||||||||||||||||||||||||||||||||||||||
Db         22 ATGRLLSASDLDLRGGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 81

Qy         53 ESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVD 112
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 ESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVD 141

Qy        113 EPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGE 172
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 EPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGE 201

Qy        173 ETELTTPVLPEETQEEDAKKTFKESREAALNLAY 206
              |||||||||||||||||||||||||||||||||
Db        202 ETELTTPVLPEETQEEDAKKTFKESREAALNLAY 235


RESULT 46
US-09-902-759-137
; Sequence 137, Application US/09902759
; Publication No. US20030077654A1
; GENERAL INFORMATION:
;  APPLICANT: Genentech, Inc.
;  APPLICANT:  Ashkenazi, Avi
;  APPLICANT:  Botstein, David
;  APPLICANT:  Desnoyers, Luc
;  APPLICANT:  Eaton, Dan L.
;  APPLICANT:  Ferrara, Napoleone
;  APPLICANT:  Filvaroff, Ellen
;  APPLICANT:  Fong, Sherman
;  APPLICANT:  Gao, Wei-Qiang
;  APPLICANT:  Gerber, Hanspeter
;  APPLICANT:  Gerritsen, Mary E.
;  APPLICANT:  Goddard, A.
;  APPLICANT:  Godowski, Paul J.
;  APPLICANT:  Grimaldi, Christopher J.
;  APPLICANT:  Gurney, Austin L.
;  APPLICANT:  Hillan, Kenneth, J.
;  APPLICANT:  Kljavin, Ivar J.
;  APPLICANT:  Mather, Jennie P.
;  APPLICANT:  Pan, James
;  APPLICANT:  Paoni, Nicholas F.
;  APPLICANT:  Roy, Margaret Ann
;  APPLICANT:  Stewart, Timothy A.
;  APPLICANT:  Tumas, Daniel
;  APPLICANT:  Williams, P. Mickey
;  APPLICANT:  Wood, William, I.
;  TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
;  TITLE OF INVENTION:  Acids Encoding the Same
;  FILE REFERENCE: 10466-14
;  CURRENT APPLICATION NUMBER: US/09/902,759
;  CURRENT FILING DATE:  2001-07-10
;  PRIOR APPLICATION NUMBER: PCT/US00/04414
;  PRIOR FILING DATE: 2000-02-22
;  PRIOR APPLICATION NUMBER: US 60/143,048
;  PRIOR FILING DATE: 1999-07-07
;  PRIOR APPLICATION NUMBER: US 60/145,698
;  PRIOR FILING DATE: 1999-07-26
;  PRIOR APPLICATION NUMBER: US 60/146,222
;  PRIOR FILING DATE: 1999-07-28
;  PRIOR APPLICATION NUMBER: PCT/US99/20594
;  PRIOR FILING DATE: 1999-09-08
;  PRIOR APPLICATION NUMBER: PCT/US99/20944
;  PRIOR FILING DATE: 1999-09-13
;  PRIOR APPLICATION NUMBER: PCT/US99/21090
;  PRIOR FILING DATE: 1999-09-15
;  PRIOR APPLICATION NUMBER: PCT/US99/21547
;  PRIOR FILING DATE: 1999-09-15
;  PRIOR APPLICATION NUMBER: PCT/US99/23089
;  PRIOR FILING DATE: 1999-10-05
;  PRIOR APPLICATION NUMBER: PCT/US99/28214
;  PRIOR FILING DATE: 1999-11-29
;  PRIOR APPLICATION NUMBER: PCT/US99/28313
;  PRIOR FILING DATE: 1999-11-30
;  PRIOR APPLICATION NUMBER: PCT/US99/28564
;  PRIOR FILING DATE: 1999-12-02
;  PRIOR APPLICATION NUMBER: PCT/US99/28565
;  PRIOR FILING DATE: 1999-12-02
;  PRIOR APPLICATION NUMBER: PCT/US99/30095
;  PRIOR FILING DATE: 1999-12-16
;  PRIOR APPLICATION NUMBER: PCT/US99/30911
;  PRIOR FILING DATE: 1999-12-20
;  PRIOR APPLICATION NUMBER: PCT/US99/30999
;  PRIOR FILING DATE: 1999-12-20
;  PRIOR APPLICATION NUMBER: PCT/US00/00219
;  PRIOR FILING DATE: 2000-01-05
;  NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 137
;   LENGTH: 382
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-902-759-137

  Query Match             98.7%;  Score 1101;  DB 10;  Length 382;
  Best Local Similarity   96.3%;  Pred. No. 8.8e-103;
  Matches  206;  Conservative    0;  Mismatches    0;  Indels    8;  Gaps    1;

Qy          1 ATGRLLS--------GQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 52
              |||||||        ||||||||||||||||||||||||||||||||||||||||||||
Db         22 ATGRLLSASDLDLRGGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 81
```

```
Qy          53 ESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVD 112
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          82 ESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVD 141

Qy         113 EPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGE 172
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         142 EPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGE 201

Qy         173 ETELTTPVLPEETQEEDAKKTFKESREAALNLAY 206
               |||||||||||||||||||||||||||||||||
Db         202 ETELTTPVLPEETQEEDAKKTFKESREAALNLAY 235


RESULT 47
US-09-902-634-137
; Sequence 137, Application US/09902634
; Publication No. US20030082540A1
; GENERAL INFORMATION:
;  APPLICANT: Genentech, Inc.
;  APPLICANT:  Ashkenazi, Avi
;  APPLICANT:  Botstein, David
;  APPLICANT:  Desnoyers, Luc
;  APPLICANT:  Eaton, Dan L.
;  APPLICANT:  Ferrara, Napoleone
;  APPLICANT:  Filvaroff, Ellen
;  APPLICANT:  Fong, Sherman
;  APPLICANT:  Gao, Wei-Qiang
;  APPLICANT:  Gerber, Hanspeter
;  APPLICANT:  Gerritsen, Mary E.
;  APPLICANT:  Goddard, A.
;  APPLICANT:  Godowski, Paul J.
;  APPLICANT:  Grimaldi, Christopher J.
;  APPLICANT:  Gurney, Austin L.
;  APPLICANT:  Hillan, Kenneth, J.
;  APPLICANT:  Kljavin, Ivar J.
;  APPLICANT:  Mather, Jennie P.
;  APPLICANT:  Pan, James
;  APPLICANT:  Paoni, Nicholas F.
;  APPLICANT:  Roy, Margaret Ann
;  APPLICANT:  Stewart, Timothy A.
;  APPLICANT:  Tumas, Daniel
;  APPLICANT:  Williams, P. Mickey
;  APPLICANT:  Wood, William, I.
;  TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
;  TITLE OF INVENTION:  Acids Encoding the Same
;  FILE REFERENCE: 10466-14
;  CURRENT APPLICATION NUMBER: US/09/902,634
;  CURRENT FILING DATE:  2001-07-10
;  PRIOR APPLICATION NUMBER: US/09/665,350
;  PRIOR FILING DATE: 2000-09-18
;  PRIOR APPLICATION NUMBER: US 60/143,048
;  PRIOR FILING DATE: 1999-07-07
;  PRIOR APPLICATION NUMBER: US 60/145,698
;  PRIOR FILING DATE: 1999-07-26
;  PRIOR APPLICATION NUMBER: US 60/146,222
;  PRIOR FILING DATE: 1999-07-28
;  PRIOR APPLICATION NUMBER: PCT/US99/20594
;  PRIOR FILING DATE: 1999-09-08
;  PRIOR APPLICATION NUMBER: PCT/US99/20944
;  PRIOR FILING DATE: 1999-09-13
;  PRIOR APPLICATION NUMBER: PCT/US99/21090
;  PRIOR FILING DATE: 1999-09-15
;  PRIOR APPLICATION NUMBER: PCT/US99/21547
;  PRIOR FILING DATE: 1999-09-15
;  PRIOR APPLICATION NUMBER: PCT/US99/23089
;  PRIOR FILING DATE: 1999-10-05
;  PRIOR APPLICATION NUMBER: PCT/US99/28214
;  PRIOR FILING DATE: 1999-11-29
;  PRIOR APPLICATION NUMBER: PCT/US99/28313
;  PRIOR FILING DATE: 1999-11-30
;  PRIOR APPLICATION NUMBER: PCT/US99/28564
;  PRIOR FILING DATE: 1999-12-02
;  PRIOR APPLICATION NUMBER: PCT/US99/28565
;  PRIOR FILING DATE: 1999-12-02
;  PRIOR APPLICATION NUMBER: PCT/US99/30095
;  PRIOR FILING DATE: 1999-12-16
;  PRIOR APPLICATION NUMBER: PCT/US99/30911
;  PRIOR FILING DATE: 1999-12-20
;  PRIOR APPLICATION NUMBER: PCT/US99/30999
;  PRIOR FILING DATE: 1999-12-20
;  PRIOR APPLICATION NUMBER: PCT/US00/00219
;  PRIOR FILING DATE: 2000-01-05
;  NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 137
;   LENGTH: 382
;   TYPE: PRT
;   ORGANISM: Homo Sapien
US-09-902-634-137

  Query Match             98.7%;  Score 1101;  DB 10;  Length 382;
  Best Local Similarity   96.3%;  Pred. No. 8.8e-103;
  Matches  206;  Conservative    0;  Mismatches    0;  Indels    8;  Gaps    1;

Qy           1 ATGRLLS--------GQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 52
               |||||||        ||||||||||||||||||||||||||||||||||||||||||||
Db          22 ATGRLLSASDLDLRGGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 81

Qy          53 ESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVD 112
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          82 ESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVD 141

Qy         113 EPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGE 172
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         142 EPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGE 201

Qy         173 ETELTTPVLPEETQEEDAKKTFKESREAALNLAY 206
               |||||||||||||||||||||||||||||||||
Db         202 ETELTTPVLPEETQEEDAKKTFKESREAALNLAY 235


RESULT 48
US-09-902-713-137
; Sequence 137, Application US/09902713
; Publication No. US20030082541A1
; GENERAL INFORMATION:
;  APPLICANT: Genentech, Inc.
;  APPLICANT:  Ashkenazi, Avi
;  APPLICANT:  Botstein, David
;  APPLICANT:  Desnoyers, Luc
;  APPLICANT:  Eaton, Dan L.
;  APPLICANT:  Ferrara, Napoleone
;  APPLICANT:  Filvaroff, Ellen
;  APPLICANT:  Fong, Sherman
;  APPLICANT:  Gao, Wei-Qiang
;  APPLICANT:  Gerber, Hanspeter
;  APPLICANT:  Gerritsen, Mary E.
;  APPLICANT:  Goddard, A.
;  APPLICANT:  Godowski, Paul J.
;  APPLICANT:  Grimaldi, Christopher J.
;  APPLICANT:  Gurney, Austin L.
;  APPLICANT:  Hillan, Kenneth, J.
;  APPLICANT:  Kljavin, Ivar J.
;  APPLICANT:  Mather, Jennie P.
;  APPLICANT:  Pan, James
;  APPLICANT:  Paoni, Nicholas F.
;  APPLICANT:  Roy, Margaret Ann
;  APPLICANT:  Stewart, Timothy A.
;  APPLICANT:  Tumas, Daniel
;  APPLICANT:  Williams, P. Mickey
;  APPLICANT:  Wood, William, I.
;  TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
;  TITLE OF INVENTION:  Acids Encoding the Same
;  FILE REFERENCE: 10466-14
;  CURRENT APPLICATION NUMBER: US/09/902,713
;  CURRENT FILING DATE:  2001-07-10
;  PRIOR APPLICATION NUMBER: 09/665,350
```

```
;  PRIOR FILING DATE: 2000-09-18
;  PRIOR APPLICATION NUMBER: PCT/US00/04414
;  PRIOR FILING DATE: 2000-02-22
;  PRIOR APPLICATION NUMBER: US 60/143,048
;  PRIOR FILING DATE: 1999-07-07
;  PRIOR APPLICATION NUMBER: US 60/145,698
;  PRIOR FILING DATE: 1999-07-26
;  PRIOR APPLICATION NUMBER: US 60/146,222
;  PRIOR FILING DATE: 1999-07-28
;  PRIOR APPLICATION NUMBER: PCT/US99/20594
;  PRIOR FILING DATE: 1999-09-08
;  PRIOR APPLICATION NUMBER: PCT/US99/20944
;  PRIOR FILING DATE: 1999-09-13
;  PRIOR APPLICATION NUMBER: PCT/US99/21090
;  PRIOR FILING DATE: 1999-09-15
;  PRIOR APPLICATION NUMBER: PCT/US99/21547
;  PRIOR FILING DATE: 1999-09-15
;  PRIOR APPLICATION NUMBER: PCT/US99/23089
;  PRIOR FILING DATE: 1999-10-05
;  PRIOR APPLICATION NUMBER: PCT/US99/28214
;  PRIOR FILING DATE: 1999-11-29
;  PRIOR APPLICATION NUMBER: PCT/US99/28313
;  PRIOR FILING DATE: 1999-11-30
;  PRIOR APPLICATION NUMBER: PCT/US99/28564
;  PRIOR FILING DATE: 1999-12-02
;  PRIOR APPLICATION NUMBER: PCT/US99/28565
;  PRIOR FILING DATE: 1999-12-02
;  PRIOR APPLICATION NUMBER: PCT/US99/30095
;  PRIOR FILING DATE: 1999-12-16
;  PRIOR APPLICATION NUMBER: PCT/US99/30911
;  PRIOR FILING DATE: 1999-12-20
;  PRIOR APPLICATION NUMBER: PCT/US99/30999
;  PRIOR FILING DATE: 1999-12-20
;  PRIOR APPLICATION NUMBER: PCT/US00/00219
;  PRIOR FILING DATE: 2000-01-05
;  NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 137
;   LENGTH: 382
;   TYPE: PRT
;   ORGANISM: Homo Sapien
US-09-902-713-137

  Query Match             98.7%;  Score 1101;  DB 10;  Length 382;
  Best Local Similarity   96.3%;  Pred. No. 8.8e-103;
  Matches  206;  Conservative    0;  Mismatches    0;  Indels    8;  Gaps    1;

Qy          1 ATGRLLS--------GQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 52
              |||||||        ||||||||||||||||||||||||||||||||||||||||||||
Db         22 ATGRLLSASDLDLRGGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 81

Qy         53 ESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVD 112
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 ESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVD 141

Qy        113 EPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGE 172
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 EPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGE 201

Qy        173 ETELTTPVLPEETQEEDAKKTFKESREAALNLAY 206
              |||||||||||||||||||||||||||||||||
Db        202 ETELTTPVLPEETQEEDAKKTFKESREAALNLAY 235


RESULT 49
US-09-907-979-137
; Sequence 137, Application US/09907979
; Publication No. US20030082542A1
; GENERAL INFORMATION:
;  APPLICANT: Genentech, Inc.
;  APPLICANT:  Ashkenazi, Avi
;  APPLICANT:  Botstein, David
;  APPLICANT:  Desnoyers, Luc
;  APPLICANT:  Eaton, Dan L.
;  APPLICANT:  Ferrara, Napoleone
;  APPLICANT:  Filvaroff, Ellen
;  APPLICANT:  Fong, Sherman
;  APPLICANT:  Gao, Wei-Qiang
;  APPLICANT:  Gerber, Hanspeter
;  APPLICANT:  Gerritsen, Mary E.
;  APPLICANT:  Goddard, A.
;  APPLICANT:  Godowski, Paul J.
;  APPLICANT:  Grimaldi, Christopher J.
;  APPLICANT:  Gurney, Austin L.
;  APPLICANT:  Hillan, Kenneth, J.
;  APPLICANT:  Kljavin, Ivar J.
;  APPLICANT:  Mather, Jennie P.
;  APPLICANT:  Pan, James
;  APPLICANT:  Paoni, Nicholas F.
;  APPLICANT:  Roy, Margaret Ann
;  APPLICANT:  Stewart, Timothy A.
;  APPLICANT:  Tumas, Daniel
;  APPLICANT:  Williams, P. Mickey
;  APPLICANT:  Wood, William, I.
;  TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
;  TITLE OF INVENTION:  Acids Encoding the Same
;  FILE REFERENCE: 10466-14
;  CURRENT APPLICATION NUMBER: US/09/907,979
;  CURRENT FILING DATE:  2001-07-17
;  PRIOR APPLICATION NUMBER: PCT/US00/04414
;  PRIOR FILING DATE: 2000-02-22
;  PRIOR APPLICATION NUMBER: US 60/143,048
;  PRIOR FILING DATE: 1999-07-07
;  PRIOR APPLICATION NUMBER: US 60/145,698
;  PRIOR FILING DATE: 1999-07-26
;  PRIOR APPLICATION NUMBER: US 60/146,222
;  PRIOR FILING DATE: 1999-07-28
;  PRIOR APPLICATION NUMBER: PCT/US99/20594
;  PRIOR FILING DATE: 1999-09-08
;  PRIOR APPLICATION NUMBER: PCT/US99/20944
;  PRIOR FILING DATE: 1999-09-13
;  PRIOR APPLICATION NUMBER: PCT/US99/21090
;  PRIOR FILING DATE: 1999-09-15
;  PRIOR APPLICATION NUMBER: PCT/US99/21547
;  PRIOR FILING DATE: 1999-09-15
;  PRIOR APPLICATION NUMBER: PCT/US99/23089
;  PRIOR FILING DATE: 1999-10-05
;  PRIOR APPLICATION NUMBER: PCT/US99/28214
;  PRIOR FILING DATE: 1999-11-29
;  PRIOR APPLICATION NUMBER: PCT/US99/28313
;  PRIOR FILING DATE: 1999-11-30
;  PRIOR APPLICATION NUMBER: PCT/US99/28564
;  PRIOR FILING DATE: 1999-12-02
;  PRIOR APPLICATION NUMBER: PCT/US99/28565
;  PRIOR FILING DATE: 1999-12-02
;  PRIOR APPLICATION NUMBER: PCT/US99/30095
;  PRIOR FILING DATE: 1999-12-16
;  PRIOR APPLICATION NUMBER: PCT/US99/30911
;  PRIOR FILING DATE: 1999-12-20
;  PRIOR APPLICATION NUMBER: PCT/US99/30999
;  PRIOR FILING DATE: 1999-12-20
;  PRIOR APPLICATION NUMBER: PCT/US00/00219
;  PRIOR FILING DATE: 2000-01-05
;  NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 137
;   LENGTH: 382
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-907-979-137

  Query Match             98.7%;  Score 1101;  DB 10;  Length 382;
  Best Local Similarity   96.3%;  Pred. No. 8.8e-103;
  Matches  206;  Conservative    0;  Mismatches    0;  Indels    8;  Gaps    1;

Qy          1 ATGRLLS--------GQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 52
              |||||||        ||||||||||||||||||||||||||||||||||||||||||||
Db         22 ATGRLLSASDLDLRGGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 81
```

```
Qy          53 ESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVD 112
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          82 ESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVD 141

Qy         113 EPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGE 172
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         142 EPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGE 201

Qy         173 ETELTTPVLPEETQEEDAKKTFKESREAALNLAY 206
               |||||||||||||||||||||||||||||||||
Db         202 ETELTTPVLPEETQEEDAKKTFKESREAALNLAY 235
```

RESULT 50
US-09-902-615-137
; Sequence 137, Application US/09902615
; Publication No. US20030092002A1
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/902,615
; CURRENT FILING DATE: 2001-12-14
; Prior application data removed. Check file wrapper or PALM.
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 137
; LENGTH: 382
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-902-615-137

Query Match 98.7%; Score 1101; DB 10; Length 382;
Best Local Similarity 96.3%; Pred. No. 8.8e-103;
Matches 206; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

```
Qy           1 ATGRLLS--------GQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 52
               |||||||        ||||||||||||||||||||||||||||||||||||||||||||
Db          22 ATGRLLSASDLDLRGGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 81

Qy          53 ESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVD 112
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          82 ESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVD 141

Qy         113 EPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGE 172
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         142 EPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGE 201

Qy         173 ETELTTPVLPEETQEEDAKKTFKESREAALNLAY 206
               |||||||||||||||||||||||||||||||||
Db         202 ETELTTPVLPEETQEEDAKKTFKESREAALNLAY 235
```

Search completed: September 9, 2004, 22:50:39
Job time : 130 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: September 9, 2004, 22:30:53 ; Search time 116 Seconds
(without alignments)
560.317 Million cell updates/sec

Title: US-09-887-855-2_COPY_22_227
Perfect score: 1115
Sequence: 1 ATGRLLSGQPVCRGGTQRPC.....EEDAKKTFKESREAALNLAY 206

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 50 summaries

Database : SPTREMBL_25:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1115	100.0	374	4	Q8TAY8	Q8tay8 homo sapien
2	1115	100.0	374	4	Q96NF3	Q96nf3 homo sapien
3	1111	99.6	374	4	Q96NC5	Q96nc5 homo sapien
4	945.5	84.8	374	11	Q9Z209	Q9z209 cricetulus
5	803.5	72.1	211	11	Q8C351	Q8c351 mus musculu
6	580	52.0	246	11	Q8BMI7	Q8bmi7 mus musculu
7	580	52.0	292	11	Q8BVU2	Q8bvu2 mus musculu
8	513	46.0	236	4	Q7Z798	Q7z798 homo sapien
9	504.5	45.2	232	4	Q7Z799	Q7z799 homo sapien
10	500	44.8	236	4	Q7Z7A0	Q7z7a0 homo sapien
11	185	16.6	1290	13	Q9W6E1	Q9w6e1 gallus gall
12	178.5	16.0	1456	11	Q61830	Q61830 mus musculu
13	177.5	15.9	1348	5	Q25199	Q25199 hydra atten
14	176.5	15.8	1479	4	Q9Y5P9	Q9y5p9 homo sapien
15	176.5	15.8	1479	4	Q9UBG0	Q9ubg0 homo sapien
16	170.5	15.3	1479	11	Q64449	Q64449 mus musculu
17	169	15.2	217	11	Q8C4F8	Q8c4f8 mus musculu
18	165.5	14.8	742	11	Q8K4Q8	Q8k4q8 mus musculu
19	165.5	14.8	742	11	Q8C979	Q8c979 mus musculu
20	161.5	14.5	134	5	Q9XYX3	Q9xyx3 hydra magni
21	159.5	14.3	742	11	Q8VIF6	Q8vif6 mus musculu
22	158	14.2	142	11	Q8CJ86	Q8cj86 mus musculu
23	158	14.2	142	11	Q8BHK7	Q8bhk7 mus musculu
24	158	14.2	295	11	Q91ZW4	Q91zw4 mus musculu
25	158	14.2	311	11	Q9D8V4	Q9d8v4 mus musculu
26	158	14.2	325	11	Q91ZX0	Q91zx0 mus musculu
27	157	14.1	158	13	Q90WI7	Q90wi7 bungarus fa
28	156	14.0	323	11	Q8CJ91	Q8cj91 mus musculu
29	156	14.0	339	6	Q95244	Q95244 sus scrofa
30	155	13.9	293	11	Q8BGZ0	Q8bgz0 mus musculu
31	155	13.9	323	11	Q8CJ94	Q8cj94 mus musculu
32	155	13.9	323	11	Q8CJ93	Q8cj93 mus musculu
33	155	13.9	323	11	Q8CJ88	Q8cj88 mus musculu
34	155	13.9	379	11	Q7TMA7	Q7tma7 mus musculu
35	155	13.9	473	11	Q7TSP9	Q7tsp9 mus musculu
36	155	13.9	477	11	Q7TSQ7	Q7tsq7 mus musculu
37	155	13.9	504	11	Q7TSQ0	Q7tsq0 mus musculu
38	155	13.9	534	11	Q7TSQ1	Q7tsq1 mus musculu
39	155	13.9	1152	13	Q90WM2	Q90wm2 xenopus lae
40	154	13.8	322	11	Q8CJ89	Q8cj89 mus musculu
41	154	13.8	323	11	Q8CJ92	Q8cj92 mus musculu
42	153	13.7	446	4	Q7Z5K9	Q7z5k9 homo sapien
43	152.5	13.7	485	6	Q95LG3	Q95lg3 odocoileus
44	152	13.6	158	13	Q90WI6	Q90wi6 bungarus mu
45	151.5	13.6	652	4	Q8IXK1	Q8ixk1 homo sapien
46	150.5	13.5	399	6	Q8HY12	Q8hy12 hylobates l
47	150	13.5	197	6	Q28008	Q28008 bos taurus
48	149.5	13.4	158	11	Q8JZX6	Q8jzx6 mus musculu
49	149.5	13.4	459	5	Q22136	Q22136 caenorhabdi
50	149	13.4	158	13	Q90WI8	Q90wi8 bungarus fa

ALIGNMENTS

RESULT 1
Q8TAY8
ID Q8TAY8 PRELIMINARY; PRT; 374 AA.
AC Q8TAY8;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Similar to unnamed protein product.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Strausberg R.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DDBJ databases.
DR EMBL; BC025407; AAH25407.1; -.
DR GO; GO:0005529; F:sugar binding; IEA.
DR InterPro; IPR001304; Lectin_C.
DR Pfam; PF00059; lectin_c; 1.
DR SMART; SM00034; CLECT; 1.
DR PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
SQ SEQUENCE 374 AA; 42312 MW; FC214E6BC9E578D9 CRC64;

Query Match 100.0%; Score 1115; DB 4; Length 374;
Best Local Similarity 100.0%; Pred. No. 7.1e-99;
Matches 206; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 ATGRLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKL 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ATGRLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKL 81

Qy         61 IEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEPSCGSEV 120
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          82 IEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEPSCGSEV 141

Qy         121 CVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGEETELTTPV 180
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         142 CVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGEETELTTPV 201

Qy         181 LPEETQEEDAKKTFKESREAALNLAY 206
                |||||||||||||||||||||||||
Db         202 LPEETQEEDAKKTFKESREAALNLAY 227


RESULT 2
Q96NF3
ID   Q96NF3       PRELIMINARY;      PRT;   374 AA.
AC   Q96NF3;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Hypothetical protein FLJ30977.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Ishibashi T., Kanehori K., Yosida M., Watanabe S., Ishida S., Ono Y.,
RA   Hotuta T., Hiraoka S., Murakawa K., Takiguchi S., Kusano J.,
RA   Watanabe M., Fujimori K., Tanai H., Ishida M., Yamashita H., Chiba Y.,
RA   Sugiyama T., Irie R., Otsuki T., Sato H., Wakamatsu A., Ishii S.,
RA   Yamamoto J., Isono Y., Kawai-Hio Y., Saito K., Nishikawa T.,
RA   Kimura K., Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K.,
RA   Wagatsuma M., Takahashi-Fujii A., Oshima A., Sugiyama A., Kawakami B.,
RA   Suzuki Y., Sugano S., Nagahari K., Masuho Y., Nagai K., Isogai T.;
RT   "NEDO human cDNA sequencing project.";
RL   Submitted (OCT-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AK055539; BAB70946.1; -.
DR   GO; GO:0005529; F:sugar binding; IEA.
DR   InterPro; IPR001304; Lectin_C.
DR   Pfam; PF00059; lectin_c; 1.
DR   SMART; SM00034; CLECT; 1.
DR   PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
KW   Hypothetical protein.
SQ   SEQUENCE   374 AA;  42280 MW;  8AE64E6BC9E56DCD CRC64;

  Query Match             100.0%;  Score 1115;  DB 4;  Length 374;
  Best Local Similarity   100.0%;  Pred. No. 7.1e-99;
  Matches  206;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy           1 ATGRLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKL 60
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          22 ATGRLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKL 81

Qy          61 IEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEPSCGSEV 120
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          82 IEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEPSCGSEV 141

Qy         121 CVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGEETELTTPV 180
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         142 CVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGEETELTTPV 201

Qy         181 LPEETQEEDAKKTFKESREAALNLAY 206
                |||||||||||||||||||||||||
Db         202 LPEETQEEDAKKTFKESREAALNLAY 227


RESULT 3
Q96NC5
ID   Q96NC5       PRELIMINARY;      PRT;   374 AA.
AC   Q96NC5;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Hypothetical protein FLJ31092.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Tashiro H., Yamazaki M., Watanabe K., Kumagai A., Itakura S.,
RA   Fukuzumi Y., Fujimori Y., Komiyama M., Sugiyama T., Irie R.,
RA   Otsuki T., Sato H., Wakamatsu A., Ishii S., Yamamoto J., Isono Y.,
RA   Kawai-Hio Y., Saito K., Nishikawa T., Kimura K., Yamashita H.,
RA   Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K., Wagatsuma M.,
RA   Murakawa K., Kanehori K., Takahashi-Fujii A., Oshima A., Sugiyama A.,
RA   Kawakami B., Suzuki Y., Sugano S., Nagahari K., Masuho Y., Nagai K.,
RA   Isogai T.;
RT   "NEDO human cDNA sequencing project.";
RL   Submitted (OCT-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AK055654; BAB70978.1; -.
DR   GO; GO:0005529; F:sugar binding; IEA.
DR   InterPro; IPR001304; Lectin_C.
DR   Pfam; PF00059; lectin_c; 1.
DR   SMART; SM00034; CLECT; 1.
DR   PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
KW   Hypothetical protein.
SQ   SEQUENCE   374 AA;  42310 MW;  CBF74E676E23BA16 CRC64;

  Query Match             99.6%;  Score 1111;  DB 4;  Length 374;
  Best Local Similarity   99.5%;  Pred. No. 1.7e-98;
  Matches  205;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy           1 ATGRLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKL 60
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          22 ATGRLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKL 81

Qy          61 IEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEPSCGSEV 120
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          82 IEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEPSCGSEV 141

Qy         121 CVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGEETELTTPV 180
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         142 CVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGEETELTTPV 201

Qy         181 LPEETQEEDAKKTFKESREAALNLAY 206
                ||||||||| ||||||||||||||||
Db         202 LPEETQEEDTKKTFKESREAALNLAY 227


RESULT 4
Q9Z209
ID   Q9Z209       PRELIMINARY;      PRT;   374 AA.
AC   Q9Z209;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Layilin.
OS   Cricetulus griseus (Chinese hamster).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC   Cricetulus.
OX   NCBI_TaxID=10029;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Borowsky M.L., Hynes R.O.;
RT   "Layilin, a novel talin-binding transmembrane protein homologous with
RT   C-type lectins, is localized in membrane ruffles.";
RL   J. Cell Biol. 143:0-0(1998).
DR   EMBL; AF093673; AAC68695.1; -.
DR   HSSP; P22897; 1EGG.
DR   GO; GO:0005529; F:sugar binding; IEA.
DR   InterPro; IPR001304; Lectin_C.
DR   Pfam; PF00059; lectin_c; 1.
DR   SMART; SM00034; CLECT; 1.
DR   PROSITE; PS50041; C_TYPE_LECTIN_2; 1
```

```
SQ   SEQUENCE   374 AA;  42435 MW;  298A8BA24FB04E1C CRC64;

  Query Match             84.8%;  Score 945.5;  DB 11;  Length 374;
  Best Local Similarity   85.9%;  Pred. No. 1.4e-82;
  Matches  176;  Conservative    9;  Mismatches   19;  Indels    1;  Gaps    1;

Qy          3 GRLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIE 62
              ||||||| ||||||:|||||||||||  :|||||||||||||||||||||:||||:|||
Db         24 GRLLSGQLVCRGGTRRPCYKVIYFHDAFQRLNFEEAKEACRRDGGQLVSIETEDEQRLIE 83

Qy         63 KFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEPSCGSEVCV 122
              ||||||| |||||||||| | || |:|||||||||||| ||||||||||||||||||||
Db         84 KFIENLLASDGDFWIGLRRLEVKQVNNTACQDLYAWTDGSTSQFRNWYVDEPSCGSEVCV 143

Qy        123 VMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPA-VPSREAEGEETELTTPVL 181
              ||||||||| |||| |||||||||||||||||||:||||:  ||    || ||  ||||
Db        144 VMYHQPSAPPGIGGSYMFQWNDDRCNMKNNFICKYADEKPSTTPSIRPGGEATEPPTPVL 203

Qy        182 PEETQEEDAKKTFKESREAALNLAY 206
              |||||:|| |:|||||||||||||
Db        204 PEETQKEDTKETFKESREAALNLAY 228


RESULT 5
Q8C351
ID   Q8C351      PRELIMINARY;      PRT;   211 AA.
AC   Q8C351;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   LAYILIN homolog (Fragment).
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Lung;
RX   MEDLINE=22354683; PubMed=12466851;
RA   The FANTOM Consortium,
RA   the RIKEN Genome Exploration Research Group Phase I & II Team;
RT   "Analysis of the mouse transcriptome based on functional annotation of
RT   60,770 full-length cDNAs.";
RL   Nature 420:563-573(2002).
DR   EMBL; AK086930; BAC39765.1; -.
DR   GO; GO:0005529; F:sugar binding; IEA.
DR   InterPro; IPR001304; Lectin_C.
DR   Pfam; PF00059; lectin_c; 1.
DR   SMART; SM00034; CLECT; 1.
DR   PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
FT   NON_TER     211     211
SQ   SEQUENCE   211 AA;  23697 MW;  AD9870B5957DD5AE CRC64;

  Query Match             72.1%;  Score 803.5;  DB 11;  Length 211;
  Best Local Similarity   79.8%;  Pred. No. 3.3e-69;
  Matches  150;  Conservative    9;  Mismatches   20;  Indels    9;  Gaps    2;

Qy          3 GRLLS--------GQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIES 54
              |||||        || ||||||:|||||||||||  :|||||||| |  |||||||||:
Db         24 GRLLSASDLDPRGGQLVCRGGTRRPCYKVIYFHDAFQRLNFEEAKETCMEDGGQLVSIET 83

Qy         55 EDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEP 114
              ||||:|||||||||| |||||||||:| ||||||:|||||||||||||| ||||||||||
Db         84 EDEQRLIEKFIENLLASDGDFWIGLKRLEEKQSNNTACQDLYAWTDGSTSQFRNWYVDEP 143

Qy        115 SCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPA-VPSREAEGEE 173
              |||||||||||||||||| |||| |||||||||||||||||:|||:  ||     ||
Db        144 SCGSEVCVVMYHQPSAPPGIGGSYMFQWNDDRCNMKNNFICKYHDDKPSTTPSIWPGGEA 203

Qy        174 TELTTPVL 181
              ||  ||:|
Db        204 TEPATPLL 211


RESULT 6
Q8BMI7
ID   Q8BMI7      PRELIMINARY;      PRT;   246 AA.
AC   Q8BMI7;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   C-type lectin protein MT75 homolog.
GN   CHODL.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Forelimb;
RX   MEDLINE=22354683; PubMed=12466851;
RA   The FANTOM Consortium,
RA   the RIKEN Genome Exploration Research Group Phase I & II Team;
RT   "Analysis of the mouse transcriptome based on functional annotation of
RT   60,770 full-length cDNAs.";
RL   Nature 420:563-573(2002).
DR   EMBL; AK031063; BAC27234.1; -.
DR   MGD; MGI:2179069; Chodl.
DR   GO; GO:0005529; F:sugar binding; IEA.
DR   InterPro; IPR001304; Lectin_C.
DR   Pfam; PF00059; lectin_c; 1.
DR   SMART; SM00034; CLECT; 1.
DR   PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
SQ   SEQUENCE   246 AA;  27394 MW;  E872660C58267752 CRC64;

  Query Match             52.0%;  Score 580;  DB 11;  Length 246;
  Best Local Similarity   57.9%;  Pred. No. 1.2e-47;
  Matches  110;  Conservative   26;  Mismatches   40;  Indels   14;  Gaps    5;

Qy          4 RLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEK 63
              |::||| ||    : ||||: |||: | |::|:||: ||  :|| |:|:|:| ||||||
Db         23 RVVSGQKVCFADVKHPCYKMAYFHELSSRVSFQEARLACESEGGVLLSLENEAEQKLIES 82

Qy         64 FIENLLP-----SDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEPSCGS 118
               ::||       |||||||| |  : |: | || ||| |:||| ||||||| |||||||
Db         83 MLQNLTKPGTGISDGDFWIGLLRSGDGQT-SGACPDLYQWSDGSSSQFRNWYTDEPSCGS 141

Qy        119 EVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDE-KPAVPSREAEGEETELT 177
              | |||||||:|  |:||||::||||||||||:|:||||  |  |  |:     |:  ||
Db        142 EKCVVMYHQPTANPGLGGPYLYQWNDDRCNMKHNYICKYEPEIHPTEPA-----EKPYLT 196

Qy        178 TPVLPEETQE 187
                 |||| |
Db        197 NQ--PEETHE 204


RESULT 7
Q8BVU2
ID   Q8BVU2      PRELIMINARY;      PRT;   292 AA.
AC   Q8BVU2;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   C-type lectin protein MT75 homolog.
GN   CHODL.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Head;
RX   MEDLINE=22354683; PubMed=12466851;
RA   The FANTOM Consortium,
RA   the RIKEN Genome Exploration Research Group Phase I & II Team;
RT   "Analysis of the mouse transcriptome based on functional annotation of
```

```
RT   60,770 full-length cDNAs.";
RL   Nature 420:563-573(2002).
DR   EMBL; AK076523; BAC36378.1; -.
DR   MGD; MGI:2179069; Chodl.
DR   GO; GO:0005529; F:sugar binding; IEA.
DR   InterPro; IPR001304; Lectin_C.
DR   Pfam; PF00059; lectin_c; 1.
DR   SMART; SM00034; CLECT; 1.
DR   PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
SQ   SEQUENCE   292 AA;  32502 MW;  73E631C0714D54E2 CRC64;

  Query Match             52.0%;  Score 580;  DB 11;  Length 292;
  Best Local Similarity   57.9%;  Pred. No. 1.5e-47;
  Matches  110;  Conservative   26;  Mismatches   40;  Indels   14;  Gaps    5;

Qy          4 RLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEK 63
              |::||| ||    : ||||: |||: | |::|:||: ||  :|| |:|:|:| ||||||
Db         23 RVVSGQKVCFADVKHPCYKMAYFHELSSRVSFQEARLACESEGGVLLSLENEAEQKLIES 82

Qy         64 FIENLLP-----SDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEPSCGS 118
               ::||       |||||||||  |  : |: | || ||| |:||| ||||||| |||||||
Db         83 MLQNLTKPGTGISDGDFWIGLLRSGDGQT-SGACPDLYQWSDGSSSQFRNWYTDEPSCGS 141

Qy        119 EVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDE-KPAVPSREAEGEETELT 177
              | |||||||:|  |:||||::||||||||||:|:||||  |  |  |:     |:  ||
Db        142 EKCVVMYHQPTANPGLGGPYLYQWNDDRCNMKHNYICKYEPEIHPTEPA-----EKPYLT 196

Qy        178 TPVLPEETQE 187
                 |||| |
Db        197 NQ--PEETHE 204


RESULT 8
Q7Z798
ID   Q7Z798      PRELIMINARY;      PRT;   236 AA.
AC   Q7Z798;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Chondrolectin variant CHODLFdeltaE.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22638322; PubMed=12621022;
RA   Weng L., Van Bockstaele D.R., Wauters J., Van Marck E., Plum J.,
RA   Berneman Z.N., Merregaert J.;
RT   "A Novel Alternative Spliced Chondrolectin Isoform Lacking the
RT   Transmembrane Domain Is Expressed during T Cell Maturation.";
RL   J. Biol. Chem. 278:19164-19170(2003).
DR   EMBL; AF523315; AAP43904.1; -.
SQ   SEQUENCE   236 AA;  27185 MW;  5110CDC80C7C2CD9 CRC64;

  Query Match             46.0%;  Score 513;  DB 4;  Length 236;
  Best Local Similarity   54.3%;  Pred. No. 3.3e-41;
  Matches   94;  Conservative   28;  Mismatches   37;  Indels   14;  Gaps    3;

Qy         25 YFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLP-----SDGDFWIGL 79
              |||: | |::|:||: ||  :|| |:|:|:| ||||||  ::||       |||||||||
Db          3 YFHELSSRVSFQEARLACESEGGVLLSLENEAEQKLIESMLQNLTKPGTGISDGDFWIGL 62

Qy         80 RRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEPSCGSEVCVVMYHQPSAPAGIGGPYM 139
                 |  : |: | || ||| |:||| ||:|||| |||||||| ||||||||:|  |:||||:
Db         63 WRNGDGQT-SGACPDLYQWSDGSNSQYRNWYTDEPSCGSEKCVVMYHQPTANPGLGGPYL 121

Qy        140 FQWNDDRCNMKNNFICKYSD--------EKPAVPSREAEGEETELTTPVLPEE 184
              :||||||||||:|:||||          ||| : ::  :  :  : |  : ||
Db        122 YQWNDDRCNMKHNYICKYEPEINPTAPVEKPYLTNQPGDTHQNVVVTEAVKEE 174


RESULT 9
```

```
Q7Z799
ID   Q7Z799      PRELIMINARY;      PRT;   232 AA.
AC   Q7Z799;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Chondrolectin variant CHODLF.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22638322; PubMed=12621022;
RA   Weng L., Van Bockstaele D.R., Wauters J., Van Marck E., Plum J.,
RA   Berneman Z.N., Merregaert J.;
RT   "A Novel Alternative Spliced Chondrolectin Isoform Lacking the
RT   Transmembrane Domain Is Expressed during T Cell Maturation.";
RL   J. Biol. Chem. 278:19164-19170(2003).
DR   EMBL; AF523314; AAP43903.1; -.
SQ   SEQUENCE   232 AA;  25979 MW;  98B80BBDF7C7F439 CRC64;

  Query Match             45.2%;  Score 504.5;  DB 4;  Length 232;
  Best Local Similarity   61.9%;  Pred. No. 2.1e-40;
  Matches   91;  Conservative   21;  Mismatches   28;  Indels    7;  Gaps    3;

Qy         25 YFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLP-----SDGDFWIGL 79
              |||: | |::|:||: ||  :|| |:|:|:| ||||||  ::||       |||||||||
Db          3 YFHELSSRVSFQEARLACESEGGVLLSLENEAEQKLIESMLQNLTKPGTGISDGDFWIGL 62

Qy         80 RRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEPSCGSEVCVVMYHQPSAPAGIGGPYM 139
                 |  : |: | || ||| |:||| ||:|||| |||||||| ||||||||:|  |:||||:
Db         63 WRNGDGQT-SGACPDLYQWSDGSNSQYRNWYTDEPSCGSEKCVVMYHQPTANPGLGGPYL 121

Qy        140 FQWNDDRCNMKNNFICKYSDE-KPAVP 165
              :||||||||||:|:||||  |  |  |
Db        122 YQWNDDRCNMKHNYICKYEPEINPTAP 148


RESULT 10
Q7Z7A0
ID   Q7Z7A0      PRELIMINARY;      PRT;   236 AA.
AC   Q7Z7A0;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Chondrolectin variant CHODLdeltaE.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22638322; PubMed=12621022;
RA   Weng L., Van Bockstaele D.R., Wauters J., Van Marck E., Plum J.,
RA   Berneman Z.N., Merregaert J.;
RT   "A Novel Alternative Spliced Chondrolectin Isoform Lacking the
RT   Transmembrane Domain Is Expressed during T Cell Maturation.";
RL   J. Biol. Chem. 278:19164-19170(2003).
DR   EMBL; AF523313; AAP43902.1; -.
SQ   SEQUENCE   236 AA;  27115 MW;  9C6CE9A9EE7062E2 CRC64;

  Query Match             44.8%;  Score 500;  DB 4;  Length 236;
  Best Local Similarity   53.2%;  Pred. No. 5.8e-40;
  Matches   92;  Conservative   28;  Mismatches   39;  Indels   14;  Gaps    3;

Qy         25 YFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLP-----SDGDFWIGL 79
              |||: | |::|: |: ||  :|| |:|:|:| ||||||  ::||       |||||||||
Db          3 YFHELSSRVSFQGARLACESEGGVLLSLENEAEQKLIESMLQNLTKPGTGISDGDFWIGL 62

Qy         80 RRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEPSCGSEVCVVMYHQPSAPAGIGGPYM 139
                 |  : |: | || ||| |:||| ||:|||| |||||||| ||||||||:|  |:| ||:
Db         63 WRNGDGQT-SGACPDLYQWSDGSNSQYRNWYTDEPSCGSEKCVVMYHQPTANPGLGSPYL 121
```

```
Qy        140 FQWNDDRCNMKNNFICKYSD--------EKPAVPSREAEGEETELTTPVLPEE 184
              :||||||||||:|:||||          ||| : ::  :  :  : |  : ||
Db        122 YQWNDDRCNMKHNYICKYEPEINPTAPVEKPYLTNQPGDTHQNVVVTEAVKEE 174
```

RESULT 11
Q9W6E1
ID Q9W6E1 PRELIMINARY; PRT; 1290 AA.
AC Q9W6E1;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Neurocan core protein.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20309833; PubMed=10851024;
RA Li H., Leung T.C., Hoffman S., Balsamo J., Lilien J.;
RT "Coordinate Regulation of Cadherin and Integrin Function by the
RT Chondroitin Sulfate Proteoglycan Neurocan.";
RL J. Cell Biol. 149:1275-1288(2000).
DR EMBL; AF116856; AAD24546.2; -.
DR HSSP; P08709; 1BF9.
DR GO; GO:0005509; F:calcium ion binding; IEA.
DR GO; GO:0005540; F:hyaluronic acid binding; IEA.
DR GO; GO:0005529; F:sugar binding; IEA.
DR GO; GO:0007155; P:cell adhesion; IEA.
DR InterPro; IPR000152; Asx_hydroxyl_S.
DR InterPro; IPR000742; EGF_2.
DR InterPro; IPR001881; EGF_Ca.
DR InterPro; IPR001438; EGF_II.
DR InterPro; IPR006209; EGF_like.
DR InterPro; IPR003599; Ig.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR001304; Lectin_C.
DR InterPro; IPR000538; Link.
DR InterPro; IPR000436; Sushi_SCR_CCP.
DR Pfam; PF00008; EGF; 2.
DR Pfam; PF00047; ig; 1.
DR Pfam; PF00059; lectin_c; 1.
DR Pfam; PF00084; sushi; 1.
DR Pfam; PF00193; Xlink; 2.
DR PRINTS; PR00010; EGFBLOOD.
DR PRINTS; PR01265; LINKMODULE.
DR ProDom; PD000918; Link; 2.
DR SMART; SM00032; CCP; 1.
DR SMART; SM00034; CLECT; 1.
DR SMART; SM00179; EGF_CA; 1.
DR SMART; SM00409; IG; 1.
DR SMART; SM00445; LINK; 2.
DR PROSITE; PS00010; ASX_HYDROXYL; 1.
DR PROSITE; PS00615; C_TYPE_LECTIN_1; 1.
DR PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
DR PROSITE; PS00022; EGF_1; 2.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS01187; EGF_CA; 1.
DR PROSITE; PS50835; IG_LIKE; 1.
DR PROSITE; PS01241; LINK; 2.
KW EGF-like domain.
SQ SEQUENCE 1290 AA; 138877 MW; 182BD86D0E40BE78 CRC64;

Query Match 16.6%; Score 185; DB 13; Length 1290;
Best Local Similarity 34.7%; Pred. No. 8.9e-09;
Matches 50; Conservative 18; Mismatches 42; Indels 34; Gaps 8;

```
Qy         17 QRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFW 76
              |  ||:  ||   ||| ::|:|:  |||  | | || |::|   |  |        : |
Db       1064 QGHCYR--YF---SRRRSWEDAERDCRRRAGHLTSIHSQEEHGFINSF------GHENTW 1112
```

```
Qy         77 IGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEPS---CGSEVCVVMY-HQPSAPA 132
              |||  |  :|        : ||| :  |: ||  ::|     | | |||:  |:
Db       1113 IGLNDRIVEQD--------FQWTDNTGLQYENWRENQPDNFFAGGEDCVVLVSHE----- 1159

Qy        133 GIGGPYMFQWNDDRCNMKNNFICK 156
                ||     :|||  ||    :|||
Db       1160 -IG-----KWNDVPCNYNLPYICK 1177
```

RESULT 12
Q61830
ID Q61830 PRELIMINARY; PRT; 1456 AA.
AC Q61830;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Macrophage mannose receptor precursor.
GN MRC1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6;
RX MEDLINE=93043353; PubMed=1421407;
RA Harris N., Rits M., Chang G., Ezekowitz R.B.;
RT "Characterization of the murine macrophage mannose receptor.";
RL Blood 80:2363-2373(1992).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6;
RA Super M.;
RL Submitted (APR-1992) to the EMBL/GenBank/DDBJ databases.
DR EMBL; Z11974; CAA78028.1; -.
DR PIR; A48925; A48925.
DR PDB; 1DQG; 10-MAY-00.
DR PDB; 1DQO; 10-MAY-00.
DR PDB; 1FWU; 17-JAN-01.
DR PDB; 1FWV; 17-JAN-01.
DR MGD; MGI:97142; Mrc1.
DR GO; GO:0009928; C:cell surface (sensu Magnoliophyta); IDA.
DR GO; GO:0005537; F:mannose binding; IDA.
DR GO; GO:0004888; F:transmembrane receptor activity; IDA.
DR InterPro; IPR002353; AntifreezeII.
DR InterPro; IPR000562; FN_Type_II.
DR InterPro; IPR001304; Lectin_C.
DR InterPro; IPR008997; RicinB_like.
DR InterPro; IPR000772; Ricin_B_lectin.
DR Pfam; PF00040; fn2; 1.
DR Pfam; PF00059; lectin_c; 8.
DR Pfam; PF00652; Ricin_B_lectin; 2.
DR PRINTS; PR00356; ANTIFREEZEII.
DR PRINTS; PR00013; FNTYPEII.
DR ProDom; PD000995; FN_Type_II; 1.
DR SMART; SM00034; CLECT; 8.
DR SMART; SM00059; FN2; 1.
DR SMART; SM00458; RICIN; 1.
DR PROSITE; PS00615; C_TYPE_LECTIN_1; 6.
DR PROSITE; PS50041; C_TYPE_LECTIN_2; 8.
DR PROSITE; PS00023; FIBRONECTIN_2; 1.
DR PROSITE; PS50231; RICIN_B_LECTIN; 1.
KW Receptor; Signal.
FT SIGNAL 1 18 POTENTIAL.
FT CHAIN 19 1456 MACROPHAGE MANNOSE RECEPTOR.
SQ SEQUENCE 1456 AA; 165065 MW; 4EBD3F1B8619A594 CRC64;

Query Match 16.0%; Score 178.5; DB 11; Length 1456;
Best Local Similarity 25.9%; Pred. No. 4.3e-08;
Matches 53; Conservative 33; Mismatches 66; Indels 53; Gaps 8;

```
Qy         21 YKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFWIGLR 80
```

```
                ||   :: :  :   : |:  |::: | | :|:|| |:| : |:| |       ::||:
Db        807 YKDYQYYFSKEKETMDNARRFCKKNFGDLATIKSESEKKFLWKYI-NKNGGQSPYFIGML 865

Qy         81 RREEKQSNSTACQDLYAWTDGSISQFRNWYVDEPSCGS--EVCVVMYHQPSAPAGIGGPY 138
              :|:         : | |||    |   |    ||:  :  | || ||      :|
Db        866 ISMDKK---------FIWMDGSKVDFVAWATGEPNFANDDENCVTMY----TNSGF---- 908

Qy        139 MFQWNDDRCNMKNNFICK---YSDEKPAVPSREAEGEETELTTPVLPEETQE-------- 187
              |||  |   |||||:    |    |:|           |||  |   :|
Db        909 ---WNDINCGYPNNFICQRHNSSINATAMP----------TTPTTPGGCKEGWHLYKNK 954

Qy        188 --------EDAKKTFKESREAALNL 204
                     : ||:::::|:|   |
Db        955 CFKIFGFANEEKKSWQDARQACKGL 979


RESULT 13
Q25199
ID   Q25199      PRELIMINARY;      PRT;  1348 AA.
AC   Q25199;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Tyrosine kinase receptor.
OS   Hydra attenuata (Hydra) (Hydra vulgaris).
OC   Eukaryota; Metazoa; Cnidaria; Hydrozoa; Hydroida; Anthomedusae;
OC   Hydridae; Hydra.
OX   NCBI_TaxID=6087;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Irvine;
RX   MEDLINE=20209407; PubMed=10744720;
RA   Reidling J.C., Miller M.A., Steele R.E.;
RT   "Sweet Tooth, a Novel Receptor Protein-tyrosine Kinase with C-type
RT   Lectin-like Extracellular Domains.";
RL   J. Biol. Chem. 275:10323-10330(2000).
DR   EMBL; L22612; AAA29218.2; -.
DR   HSSP; P11362; 1FGK.
DR   GO; GO:0005524; F:ATP binding; IEA.
DR   GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.
DR   GO; GO:0004872; F:receptor activity; IEA.
DR   GO; GO:0005529; F:sugar binding; IEA.
DR   GO; GO:0016740; F:transferase activity; IEA.
DR   GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR   InterPro; IPR001304; Lectin_C.
DR   InterPro; IPR003990; Pancreatis_ac.
DR   InterPro; IPR000719; Prot_kinase.
DR   InterPro; IPR001245; Tyr_pkinase.
DR   InterPro; IPR008266; Tyr_pkinase_AS.
DR   Pfam; PF00059; lectin_c; 4.
DR   Pfam; PF00069; pkinase; 1.
DR   PRINTS; PR01504; PNCREATITSAP.
DR   ProDom; PD000001; Prot_kinase; 2.
DR   SMART; SM00034; CLECT; 4.
DR   SMART; SM00219; TyrKc; 1.
DR   PROSITE; PS00615; C_TYPE_LECTIN_1; 2.
DR   PROSITE; PS50041; C_TYPE_LECTIN_2; 4.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR   PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
KW   ATP-binding; Kinase; Receptor; Transferase; Tyrosine-protein kinase.
SQ   SEQUENCE   1348 AA;  156916 MW;  77D2122093227FEF CRC64;

  Query Match             15.9%;  Score 177.5;  DB 5;  Length 1348;
  Best Local Similarity   28.0%;  Pred. No. 4.9e-08;
  Matches   47;  Conservative   32;  Mismatches   46;  Indels   43;  Gaps    8;

Qy          6 LSGQPVC--RGGTQRPC------YKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDE 57
              || : :|  :  |   |       |::  :    |   :: ::  :|:  || |:|||:::|
Db        412 LSHRFICKVKRATNEYCAEGWTSYRIYCYFIYSIEFDWFKSFSSCQNIGGNLLSIENQEE 471

Qy         58 QKLIEKFIENLLPSDGD-FWIGLRR------REEKQSNSTACQDLYAWTDGSISQFRNWY 110
                 :|||| |  | | :|||| :      :: |:         : |:| : :|| ||
Db        472 ----NRFIENDLIKDNDKYWIGLNKIWNDYLKKNKR---------FEWSDNTYTQFFNWI 518

Qy        111 VDEP--SCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICK 156
                  ::|  : | | || | :              |:|  | : | ||||
Db        519 TNQPDNNNGIESCVEMNYN------------GWSDKECKVLNGFICK 553


RESULT 14
Q9Y5P9
ID   Q9Y5P9      PRELIMINARY;      PRT;  1479 AA.
AC   Q9Y5P9;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Endocytic receptor Endo180.
GN   ENDO180.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20148849; PubMed=10683150;
RA   Sheikh H., Yarwood H., Ashworth A., Isacke C.;
RT   "Endo180, an endocytic recycling glycoprotein related to the
RT   macrophage mannose receptor is expressed on fibroblasts, endothelial
RT   cells and macrophages and functions as a lectin receptor.";
RL   J. Cell Sci. 113:1021-1032(2000).
DR   EMBL; AF134838; AAD30280.1; -.
DR   HSSP; P02751; 2FN2.
DR   GO; GO:0004872; F:receptor activity; IEA.
DR   GO; GO:0005529; F:sugar binding; IEA.
DR   GO; GO:0005215; F:transporter activity; IEA.
DR   GO; GO:0006810; P:transport; IEA.
DR   InterPro; IPR000562; FN_Type_II.
DR   InterPro; IPR001304; Lectin_C.
DR   InterPro; IPR000566; Lipocln_cytFABP.
DR   InterPro; IPR008997; RicinB_like.
DR   InterPro; IPR000772; Ricin_B_lectin.
DR   Pfam; PF00040; fn2; 1.
DR   Pfam; PF00059; lectin_c; 8.
DR   PRINTS; PR00013; FNTYPEII.
DR   ProDom; PD000995; FN_Type_II; 1.
DR   SMART; SM00034; CLECT; 8.
DR   SMART; SM00059; FN2; 1.
DR   SMART; SM00458; RICIN; 1.
DR   PROSITE; PS00615; C_TYPE_LECTIN_1; 3.
DR   PROSITE; PS50041; C_TYPE_LECTIN_2; 8.
DR   PROSITE; PS00023; FIBRONECTIN_2; 1.
DR   PROSITE; PS00213; LIPOCALIN; 1.
DR   PROSITE; PS50231; RICIN_B_LECTIN; 1.
KW   Receptor.
SQ   SEQUENCE   1479 AA;  166669 MW;  9F4BAF355F036FCE CRC64;

  Query Match             15.8%;  Score 176.5;  DB 4;  Length 1479;
  Best Local Similarity   30.4%;  Pred. No. 6.9e-08;
  Matches   51;  Conservative   25;  Mismatches   57;  Indels   35;  Gaps    7;

Qy          7 SGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIE 66
              | ||       |  ||::      : : :::|:|:|| | || |||| |  | : | | |:
Db        385 SWQPF-----QGHCYRL-----QAEKRSWQESKKACLRGGGDLVSIHSMAELEFITKQIK 434

Qy         67 NLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEPS---CGSEVCVV 123
                :    : ||||        |    | : |:|||:  | :|:  ||:       | ||
Db        435 QEVE---ELWIGL--------NDLKLQMNFEWSDGSLVSFTHWHPFEPNNFRDSLEDCVT 483

Qy        124 MYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEG 171
              ::     | |       :|||  ||     ||| : :     : |  |
Db        484 IW----GPEG-------RWNDSPCNQSLPSICKKAGQLSQGAAEEDHG 520


RESULT 15
Q9UBG0
```

```
ID   Q9UBG0      PRELIMINARY;      PRT;  1479 AA.
AC   Q9UBG0;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Urokinase receptor-associated protein UPARAP.
GN   KIAA0709.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Behrendt N., Jensen O.N., Engelholm L.H., Mortz E., Mann M., Dano K.;
RT   "A urokinase receptor-associated protein with specific collagen-
RT   binding properties.";
RL   Submitted (NOV-1998) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Brain;
RX   MEDLINE=98403880; PubMed=9734811;
RA   Ishikawa K., Nagase T., Suyama M., Miyajima N., Tanaka A., Kotani H.,
RA   Nomura N., Ohara O.;
RT   "Prediction of the coding sequences of unidentified human genes. X.
RT   The complete sequences of 100 new cDNA clones from brain which can
RT   code for large proteins in vitro.";
RL   DNA Res. 5:169-176(1998).
DR   EMBL; AF107292; AAF14192.1; -.
DR   EMBL; AB014609; BAA31684.1; -.
DR   HSSP; P02751; 2FN2.
DR   Genew; HGNC:16875; MRC2.
DR   GO; GO:0016301; F:kinase activity; IEA.
DR   GO; GO:0004872; F:receptor activity; IEA.
DR   GO; GO:0005529; F:sugar binding; IEA.
DR   GO; GO:0005215; F:transporter activity; IEA.
DR   GO; GO:0006810; P:transport; IEA.
DR   InterPro; IPR000562; FN_Type_II.
DR   InterPro; IPR001304; Lectin_C.
DR   InterPro; IPR000566; Lipocln_cytFABP.
DR   InterPro; IPR008997; RicinB_like.
DR   InterPro; IPR000772; Ricin_B_lectin.
DR   Pfam; PF00040; fn2; 1.
DR   Pfam; PF00059; lectin_c; 8.
DR   PRINTS; PR00013; FNTYPEII.
DR   ProDom; PD000995; FN_Type_II; 1.
DR   SMART; SM00034; CLECT; 8.
DR   SMART; SM00059; FN2; 1.
DR   SMART; SM00458; RICIN; 1.
DR   PROSITE; PS00615; C_TYPE_LECTIN_1; 3.
DR   PROSITE; PS50041; C_TYPE_LECTIN_2; 8.
DR   PROSITE; PS00023; FIBRONECTIN_2; 1.
DR   PROSITE; PS00213; LIPOCALIN; 1.
DR   PROSITE; PS50231; RICIN_B_LECTIN; 1.
KW   Kinase; Receptor.
SQ   SEQUENCE   1479 AA;  166654 MW;  C7583EA78E2792D1 CRC64;

  Query Match             15.8%;  Score 176.5;  DB 4;  Length 1479;
  Best Local Similarity   30.4%;  Pred. No. 6.9e-08;
  Matches   51;  Conservative   25;  Mismatches   57;  Indels   35;  Gaps    7;

Qy          7 SGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIE 66
              | ||       |  ||::      : : :::|:|:|| | || |||| |  | : | | |:
Db        385 SWQPF-----QGHCYRL-----QAEKRSWQESKKACLRGGGDLVSIHSMAELEFITKQIK 434

Qy         67 NLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEPS---CGSEVCVV 123
                :    : ||||         |    |  : |:|||:  | :|:  ||:      | ||
Db        435 QEVE---ELWIGL--------NDLKLQMNFEWSDGSLVSFTHWHPFEPNNFRDSLEDCVT 483

Qy        124 MYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEG 171
              ::      | |        :|||  ||     ||| : :     : |  |
Db        484 IW----GPEG-------RWNDSPCNQSLPSICKKAGQLSQGAAEEDHG 520
```

```
RESULT 16
Q64449
ID   Q64449      PRELIMINARY;      PRT;  1479 AA.
AC   Q64449;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Lectin lambda.
GN   MRC2.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=96355501; PubMed=8702911;
RA   Wu K., Yuan J., Lasky L.A.;
RT   "Characterization of a novel member of the macrophage mannose receptor
RT   type C lectin family.";
RL   J. Biol. Chem. 271:21323-21330(1996).
DR   EMBL; U56734; AAC52729.1; -.
DR   PIR; T42710; T42710.
DR   HSSP; P02751; 2FN2.
DR   MGD; MGI:107818; Mrc2.
DR   GO; GO:0005529; F:sugar binding; IEA.
DR   GO; GO:0005215; F:transporter activity; IEA.
DR   GO; GO:0006118; P:electron transport; IEA.
DR   GO; GO:0006810; P:transport; IEA.
DR   InterPro; IPR002353; AntifreezeII.
DR   InterPro; IPR001128; Cytochrome_P450.
DR   InterPro; IPR000562; FN_Type_II.
DR   InterPro; IPR001304; Lectin_C.
DR   InterPro; IPR000566; Lipocln_cytFABP.
DR   InterPro; IPR008997; RicinB_like.
DR   InterPro; IPR000772; Ricin_B_lectin.
DR   Pfam; PF00040; fn2; 1.
DR   Pfam; PF00059; lectin_c; 8.
DR   PRINTS; PR00356; ANTIFREEZEII.
DR   PRINTS; PR00013; FNTYPEII.
DR   ProDom; PD000995; FN_Type_II; 1.
DR   SMART; SM00034; CLECT; 8.
DR   SMART; SM00059; FN2; 1.
DR   SMART; SM00458; RICIN; 1.
DR   PROSITE; PS00086; CYTOCHROME_P450; 1.
DR   PROSITE; PS00615; C_TYPE_LECTIN_1; 3.
DR   PROSITE; PS50041; C_TYPE_LECTIN_2; 8.
DR   PROSITE; PS00023; FIBRONECTIN_2; 1.
DR   PROSITE; PS00213; LIPOCALIN; 1.
DR   PROSITE; PS50231; RICIN_B_LECTIN; 1.
SQ   SEQUENCE   1479 AA;  167112 MW;  62D456E10B9B48C1 CRC64;

  Query Match             15.3%;  Score 170.5;  DB 11;  Length 1479;
  Best Local Similarity   31.4%;  Pred. No. 2.6e-07;
  Matches   48;  Conservative   22;  Mismatches   48;  Indels   35;  Gaps    7;

Qy          7 SGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIE 66
              | ||       |  ||::      : : :::|:| || | || |:|| |  | : | | |:
Db        384 SWQPF-----QGHCYRL-----QAEKRSWQESKRACLRGGGDLLSIHSMAELEFITKQIK 433

Qy         67 NLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEPS---CGSEVCVV 123
                :    : ||||         |    |  : |:|||:  | :|:  ||:      | ||
Db        434 QEVE---ELWIGL--------NDLKLQMNFEWSDGSLVSFTHWHPFEPNNFRDSLEDCVT 482

Qy        124 MYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICK 156
              ::      | |        :|.||  ||     |||
Db        483 IW----GPEG-------RWNDSPCNQSLPSICK 504


RESULT 17
Q8C4F8
ID   Q8C4F8      PRELIMINARY;      PRT;   217 AA.
AC   Q8C4F8;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
```

```
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Chondroitin sulfate proteoglycan 3 (Fragment).
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Cerebellum;
RX   MEDLINE=22354683; PubMed=12466851;
RA   The FANTOM Consortium,
RA   the RIKEN Genome Exploration Research Group Phase I & II Team;
RT   "Analysis of the mouse transcriptome based on functional annotation of
RT   60,770 full-length cDNAs.";
RL   Nature 420:563-573(2002).
DR   EMBL; AK082298; BAC38458.1; -.
DR   GO; GO:0005529; F:sugar binding; IEA.
DR   InterPro; IPR002353; AntifreezeII.
DR   InterPro; IPR001304; Lectin_C.
DR   InterPro; IPR000436; Sushi_SCR_CCP.
DR   Pfam; PF00059; lectin_c; 1.
DR   Pfam; PF00084; sushi; 1.
DR   PRINTS; PR00356; ANTIFREEZEII.
DR   SMART; SM00032; CCP; 1.
DR   SMART; SM00034; CLECT; 1.
DR   PROSITE; PS00615; C_TYPE_LECTIN_1; 1.
DR   PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
FT   NON_TER       1      1
SQ   SEQUENCE   217 AA;  25949 MW;  5F1A6A57B05B76D6 CRC64;

  Query Match             15.2%;  Score 169;  DB 11;  Length 217;
  Best Local Similarity   31.1%;  Pred. No. 3.6e-08;
  Matches   42;  Conservative   16;  Mismatches   47;  Indels   30;  Gaps    5;

Qy         25 YFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFWIGLRRREE 84
              ||   : |  :|:|:  |||  | | |: | :| | |  |        : ||||  |
Db          3 YF---AHRRAWEDAERDCRRRAGHLTSVHSPEEHKFINSF------GHENSWIGLNDRTV 53

Qy         85 KQSNSTACQDLYAWTDGSISQFRNWYVDEPS---CGSEVCVVMYHQPSAPAGIGGPYMFQ 141
              ::         : ||| :  |: ||   :|     | | ||||    |           :
Db         54 ERD--------FQWTDNTGLQYENWREKQPDNFFAGGEDCVVMVAHESG----------R 95

Qy        142 WNDDRCNMKNNFICK 156
              |||  ||    ::||
Db         96 WNDVPCNYNLPYVCK 110


RESULT 18
Q8K4Q8
ID   Q8K4Q8      PRELIMINARY;      PRT;   742 AA.
AC   Q8K4Q8;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Collectin placenta 1.
GN   COLEC12 OR CL-P1.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Ohtani K., Suzuki Y., Eda S., Kawai T., Kase T., Keshi H., Sakai Y.,
RA   Fukuoh A., Sakamoto T., Itabe H., Suzutani T., Ogasawara M.,
RA   Yoshida I., Wakamiya N.;
RT   "cDNA cloning of mouse CL-P1 gene.";
RL   Submitted (JAN-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AB078434; BAC05523.1; -.
DR   MGD; MGI:2152907; Colec12.
DR   GO; GO:0006955; P:immune response; IDA.
DR   GO; GO:0006910; P:phagocytosis, binding; IDA.
DR   InterPro; IPR002353; AntifreezeII.
DR   InterPro; IPR008160; Collagen.
DR   InterPro; IPR001304; Lectin_C.
DR   Pfam; PF01391; Collagen; 2.
DR   Pfam; PF00059; lectin_c; 1.
DR   PRINTS; PR00356; ANTIFREEZEII.
DR   SMART; SM00034; CLECT; 1.
DR   PROSITE; PS00615; C_TYPE_LECTIN_1; 1.
DR   PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
SQ   SEQUENCE   742 AA;  81304 MW;  1537C490E5911C45 CRC64;

  Query Match             14.8%;  Score 165.5;  DB 11;  Length 742;
  Best Local Similarity   33.8%;  Pred. No. 3.4e-07;
  Matches   50;  Conservative   17;  Mismatches   54;  Indels   27;  Gaps    8;

Qy         20 CYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFWIGL 79
              ||   ||  :  :  ||:||  |       || | | :||: |:|           ||||
Db        618 CY---YF--SLEKEIFEDAKLFCEDKSSHLVFINSREEQQWIKKH----TVGRESHWIGL 668

Qy         80 RRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEP-SCGSEVCVVMYHQPSAPAGIGGPY 138
                |::          : | |||   ::||   :| : ||       | |      |  |
Db        669 TDSEQESE--------WKWLDGSPVDYKNWKAGQPDNWGSG------HGPGEDCA-GLIY 713

Qy        139 MFQWNDDRCNMKNNFICKYSDEKPAVPS 166
              |||| :|:  |||||:   |: ||||
Db        714 AGQWNDFQCDEINNFICE--KEREAVPS 739


RESULT 19
Q8C979
ID   Q8C979      PRELIMINARY;      PRT;   742 AA.
AC   Q8C979;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Collectin sub-family member 12.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Cerebellum;
RX   MEDLINE=22354683; PubMed=12466851;
RA   The FANTOM Consortium,
RA   the RIKEN Genome Exploration Research Group Phase I & II Team;
RT   "Analysis of the mouse transcriptome based on functional annotation of
RT   60,770 full-length cDNAs.";
RL   Nature 420:563-573(2002).
DR   EMBL; AK042772; BAC31361.1; -.
DR   GO; GO:0005529; F:sugar binding; IEA.
DR   InterPro; IPR002353; AntifreezeII.
DR   InterPro; IPR008160; Collagen.
DR   InterPro; IPR001304; Lectin_C.
DR   Pfam; PF01391; Collagen; 2.
DR   Pfam; PF00059; lectin_c; 1.
DR   PRINTS; PR00356; ANTIFREEZEII.
DR   SMART; SM00034; CLECT; 1.
DR   PROSITE; PS00615; C_TYPE_LECTIN_1; 1.
DR   PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
SQ   SEQUENCE   742 AA;  81289 MW;  1350DCCB34522C45 CRC64;

  Query Match             14.8%;  Score 165.5;  DB 11;  Length 742;
  Best Local Similarity   33.8%;  Pred. No. 3.4e-07;
  Matches   50;  Conservative   17;  Mismatches   54;  Indels   27;  Gaps    8;

Qy         20 CYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFWIGL 79
              ||   ||  :  :  ||:||  |       || | | :||: |:|           ||||
Db        618 CY---YF--SLEKEIFEDAKLFCEDKSSHLVFINSREEQQWIKKH----TVGRESHWIGL 668

Qy         80 RRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEP-SCGSEVCVVMYHQPSAPAGIGGPY 138
                |::          : | |||   ::||   :| : ||       | |      |  |
Db        669 TDSEQESE--------WKWLDGSPVDYKNWKAGQPDNWGSG------HGPGEDCA-GLIY 713
```

```
Qy        139 MFQWNDDRCNMKNNFICKYSDEKPAVPS 166
                ||||  :|:  |||||:   |: ||||
Db        714 AGQWNDFQCDEINNFICE--KEREAVPS 739
```

RESULT 20
Q9XYX3
ID Q9XYX3 PRELIMINARY; PRT; 134 AA.
AC Q9XYX3;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Receptor protein-tyrosine kinase (Fragment).
GN HTK28.
OS Hydra magnipapillata (Hydra).
OC Eukaryota; Metazoa; Cnidaria; Hydrozoa; Hydroida; Anthomedusae;
OC Hydridae; Hydra.
OX NCBI_TaxID=6085;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=105;
RX MEDLINE=20209407; PubMed=10744720;
RA Reidling J.C., Miller M.A., Steele R.E.;
RT "Sweet Tooth, a Novel Receptor Protein-tyrosine Kinase with C-type
RT Lectin-like Extracellular Domains.";
RL J. Biol. Chem. 275:10323-10330(2000).
DR EMBL; AF129528; AAD30040.1; -.
DR HSSP; P22897; 1EGG.
DR GO; GO:0016301; F:kinase activity; IEA.
DR GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.
DR GO; GO:0005529; F:sugar binding; IEA.
DR InterPro; IPR001304; Lectin_C.
DR Pfam; PF00059; lectin_c; 1.
DR SMART; SM00034; CLECT; 1.
DR PROSITE; PS00615; C_TYPE_LECTIN_1; 1.
DR PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
KW Kinase; Tyrosine-protein kinase.
FT NON_TER 1 1
FT NON_TER 134 134
SQ SEQUENCE 134 AA; 15701 MW; E7B7211C881009BC CRC64;

Query Match 14.5%; Score 161.5; DB 5; Length 134;
Best Local Similarity 28.8%; Pred. No. 1.1e-07;
Matches 40; Conservative 23; Mismatches 51; Indels 25; Gaps 5;

```
Qy         20 CYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFWIGL 79
              ||   : : | :  |: :|  :|:  || |:||| : |   |  |:::      ::||||
Db         16 CY--FFQNKTLQAKNWRDASLSCQALGGHLLSIEDQAENFFILNFLKDSSMQQDNYWIGL 73

Qy         80 RRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEPS--CGSEVCVVMYHQPSAPAGIGGP 137
                     |  :    : |:|  | || ||   :|:     : ||          :|
Db         74 --------NDASNNREFRWSDDKIPQFFNWLPKKPNNDQSEQNCV-------ETNSMG-- 116

Qy        138 YMFQWNDDRCNMKNNFICK 156
                 |||: |:  | ||||
Db        117 ----WNDENCDATNGFICK 131
```

RESULT 21
Q8VIF6
ID Q8VIF6 PRELIMINARY; PRT; 742 AA.
AC Q8VIF6;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Scavenger receptor with C-type lectin.
GN COLEC12 OR SRCL.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21575692; PubMed=11718900;
RA Nakamura K., Funakoshi H., Tokunaga F., Nakamura T.;
RT "Molecular cloning of a mouse scavenger receptor with C-type lectin
RT (SRCL)(1), a novel member of the scavenger receptor family.";
RL Biochim. Biophys. Acta 1522:53-58(2001).
DR EMBL; AB038519; BAB82497.1; -.
DR MGD; MGI:2152907; Colec12.
DR GO; GO:0006955; P:immune response; IDA.
DR GO; GO:0006910; P:phagocytosis, binding; IDA.
DR InterPro; IPR002353; AntifreezeII.
DR InterPro; IPR008160; Collagen.
DR InterPro; IPR001304; Lectin_C.
DR Pfam; PF01391; Collagen; 2.
DR Pfam; PF00059; lectin_c; 1.
DR PRINTS; PR00356; ANTIFREEZEII.
DR SMART; SM00034; CLECT; 1.
DR PROSITE; PS00615; C_TYPE_LECTIN_1; 1.
DR PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
KW Receptor; Lectin.
SQ SEQUENCE 742 AA; 81307 MW; 85A90D3AE881DB6B CRC64;

Query Match 14.3%; Score 159.5; DB 11; Length 742;
Best Local Similarity 33.1%; Pred. No. 1.3e-06;
Matches 49; Conservative 17; Mismatches 55; Indels 27; Gaps 8;

```
Qy         20 CYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFWIGL 79
              ||   ||  :  :   |:||  |      || | | :||: |:|          ||||
Db        618 CY---YF--SLEKEILEDAKLFCEDKSSHLVFINSREEQQWIKKH----TVGRESHWIGL 668

Qy         80 RRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEP-SCGSEVCVVMYHQPSAPAGIGGPY 138
                 |::          : | |||   ::||   :| : ||       | |      | |
Db        669 TDSEQESE--------WKWLDGSPVDYKNWKAGQPDNWGSG------HGPGEDCA-GLIY 713

Qy        139 MFQWNDDRCNMKNNFICKYSDEKPAVPS 166
                ||||  :|:  |||||:   |: ||||
Db        714 AGQWNDFQCDEINNFICE--KEREAVPS 739
```

RESULT 22
Q8CJ86
ID Q8CJ86 PRELIMINARY; PRT; 142 AA.
AC Q8CJ86;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE SIGNR1 (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Swiss Webster; TISSUE=Skin;
RX MEDLINE=22133304; PubMed=12137941;
RA Parent S.A., Zhang T., Chrebet G., Clemas J.A., Figueroa D.J., Ky B.,
RA Blevins R.A., Austin C.P., Rosen H.;
RT "Molecular characterization of the murine SIGNR1 gene encoding a C-
RT type lectin homologous to human DC-SIGN and DC-SIGNR.";
RL Gene 293:33-46(2002).
DR EMBL; AF424802; AAN75597.1; -.
DR GO; GO:0005529; F:sugar binding; IEA.
DR InterPro; IPR002353; AntifreezeII.
DR InterPro; IPR001304; Lectin_C.
DR Pfam; PF00059; lectin_c; 1.
DR PRINTS; PR00356; ANTIFREEZEII.
DR SMART; SM00034; CLECT; 1.
DR PROSITE; PS00615; C_TYPE_LECTIN_1; 1.
DR PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
FT NON_TER 1 1
SQ SEQUENCE 142 AA; 16437 MW; 0CBB36A383D025EF CRC64;

Query Match 14.2%; Score 158; DB 11; Length 142;
Best Local Similarity 32.4%; Pred. No. 2.5e-07;

```
  Matches   48;  Conservative   21;  Mismatches   45;  Indels   34;  Gaps    9;

Qy          25 YFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFWIGLRRREE 84
               ||    |:| |: :|  ||:    ||| | |::||     |::    : |  |:||   ::
Db          24 YFFSKSQR-NWNDAVTACKEVKAQLVIINSDEEQ----TFLQQTSKAKGPTWMGLSDLKK 78

Qy          85 KQSNSTACQDLYAWTDGSI--SQFRN-WYVDEP-SCGSEVCVVMYHQPSAPAGIGGPYMF 140
               : :        : | |||   |:|:  |   || : | | ||         || |
Db          79 EAT--------WLWVDGSTLSSRFQKYWNRGEPNNIGEEDCVEF-------AGDG----- 118

Qy         141 QWNDDRCNMKNNFICKYSDEKPAVPSRE 168
                ||| :| :|  :|||    | | |  |
Db         119 -WNDSKCELKKLWICK----KSATPCTE 141


RESULT 23
Q8BHK7
ID   Q8BHK7      PRELIMINARY;      PRT;   142 AA.
AC   Q8BHK7;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   SIGNR1 (Fragment).
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Swiss Webster; TISSUE=Spleen;
RX   MEDLINE=22133304; PubMed=12137941;
RA   Parent S.A., Zhang T., Chrebet G., Clemas J.A., Figueroa D.J., Ky B.,
RA   Blevins R.A., Austin C.P., Rosen H.;
RT   "Molecular characterization of the murine SIGNR1 gene encoding a C-
RT   type lectin homologous to human DC-SIGN and DC-SIGNR.";
RL   Gene 293:33-46(2002).
DR   EMBL; AF424800; AAN75595.1; -.
DR   EMBL; AF424801; AAN75596.1; -.
DR   GO; GO:0005529; F:sugar binding; IEA.
DR   InterPro; IPR002353; AntifreezeII.
DR   InterPro; IPR001304; Lectin_C.
DR   Pfam; PF00059; lectin_c; 1.
DR   PRINTS; PR00356; ANTIFREEZEII.
DR   SMART; SM00034; CLECT; 1.
DR   PROSITE; PS00615; C_TYPE_LECTIN_1; 1.
DR   PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
FT   NON_TER       1       1
SQ   SEQUENCE   142 AA;  16471 MW;  0CBB36A383D887EF CRC64;

  Query Match             14.2%;  Score 158;  DB 11;  Length 142;
  Best Local Similarity   32.4%;  Pred. No. 2.5e-07;
  Matches   48;  Conservative   21;  Mismatches   45;  Indels   34;  Gaps    9;

Qy          25 YFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFWIGLRRREE 84
               ||    |:| |: :|  ||:    ||| | |::||     |::    : |  |:||   ::
Db          24 YFFSKSQR-NWNDAVTACKEVKAQLVIINSDEEQ----TFLQQTSKAKGPTWMGLSDLKK 78

Qy          85 KQSNSTACQDLYAWTDGSI--SQFRN-WYVDEP-SCGSEVCVVMYHQPSAPAGIGGPYMF 140
               : :        : | |||   |:|:  |   || : | | ||         || |
Db          79 EAT--------WLWVDGSTLSSRFQKYWNRGEPNNIGEEDCVEF-------AGDG----- 118

Qy         141 QWNDDRCNMKNNFICKYSDEKPAVPSRE 168
                ||| :| :|  :|||    | | |  |
Db         119 -WNDSKCELKKFWICK----KSATPCTE 141


RESULT 24
Q91ZW4
ID   Q91ZW4      PRELIMINARY;      PRT;   295 AA.
AC   Q91ZW4;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   SIGNR1 TM-less isoform.
GN   CD209B OR SIGNR1.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6;
RA   Park C.G., Steinman R.M.;
RT   "Alternatively Spliced Forms of Mouse DC-SIGN Homologs.";
RL   Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF374471; AAL27540.1; -.
DR   MGD; MGI:1916415; Cd209b.
DR   GO; GO:0005529; F:sugar binding; IEA.
DR   InterPro; IPR002353; AntifreezeII.
DR   InterPro; IPR001304; Lectin_C.
DR   Pfam; PF00059; lectin_c; 1.
DR   PRINTS; PR00356; ANTIFREEZEII.
DR   SMART; SM00034; CLECT; 1.
DR   PROSITE; PS00615; C_TYPE_LECTIN_1; 1.
DR   PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
SQ   SEQUENCE   295 AA;  33888 MW;  A491F7D3551A91D0 CRC64;

  Query Match             14.2%;  Score 158;  DB 11;  Length 295;
  Best Local Similarity   32.4%;  Pred. No. 5.9e-07;
  Matches   48;  Conservative   21;  Mismatches   45;  Indels   34;  Gaps    9;

Qy          25 YFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFWIGLRRREE 84
               ||    |:| |: :|  ||:    ||| | |::||     |::    : |  |:||   ::
Db         177 YFFSKSQR-NWNDAVTACKEVKAQLVIINSDEEQ----TFLQQTSKAKGPTWMGLSDLKK 231

Qy          85 KQSNSTACQDLYAWTDGSI--SQFRN-WYVDEP-SCGSEVCVVMYHQPSAPAGIGGPYMF 140
               : :        : | |||   |:|:  |   || : | | ||         || |
Db         232 EAT--------WLWVDGSTLSSRFQKYWNRGEPNNIGEEDCVEF-------AGDG----- 271

Qy         141 QWNDDRCNMKNNFICKYSDEKPAVPSRE 168
                ||| :| :|  :|||    | | |  |
Db         272 -WNDSKCELKKFWICK----KSATPCTE 294


RESULT 25
Q9D8V4
ID   Q9D8V4      PRELIMINARY;      PRT;   311 AA.
AC   Q9D8V4;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   1810030I22Rik protein.
GN   CD209B OR 1810030I22RIK.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Pancreas;
RX   MEDLINE=21085660; PubMed=11217851;
RA   Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA   Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA   Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA   Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA   Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA   Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA   Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA   Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA   Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA   Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA   Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA   Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA   Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA   Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA   Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
```

```
RA   Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA   Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA   Hayashizaki Y.;
RT   "Functional annotation of a full-length mouse cDNA collection.";
RL   Nature 409:685-690(2001).
DR   EMBL; AK007656; BAB25166.1; -.
DR   HSSP; P22897; 1EGG.
DR   MGD; MGI:1916415; Cd209b.
DR   GO; GO:0005529; F:sugar binding; IEA.
DR   InterPro; IPR002353; AntifreezeII.
DR   InterPro; IPR001304; Lectin_C.
DR   Pfam; PF00059; lectin_c; 1.
DR   PRINTS; PR00356; ANTIFREEZEII.
DR   SMART; SM00034; CLECT; 1.
DR   PROSITE; PS00615; C_TYPE_LECTIN_1; 1.
DR   PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
SQ   SEQUENCE   311 AA;  35618 MW;  92353503D2EF9041 CRC64;

  Query Match             14.2%;  Score 158;  DB 11;  Length 311;
  Best Local Similarity   32.4%;  Pred. No. 6.3e-07;
  Matches   48;  Conservative   21;  Mismatches   45;  Indels   34;  Gaps    9;

Qy         25 YFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFWIGLRRREE 84
              ||    |:| |: :|  ||:    ||| | |::||     |::    : |  |:||   ::
Db        193 YFFSKSQR-NWNDAVTACKEVKAQLVIINSDEEQ----TFLQQTSKAKGPTWMGLSDLKK 247

Qy         85 KQSNSTACQDLYAWTDGSI--SQFRN-WYVDEP-SCGSEVCVVMYHQPSAPAGIGGPYMF 140
              : :        : | |||   |:|:  |   || : | | ||         || |
Db        248 EAT--------WLWVDGSTLSSRFQKYWNRGEPNNIGEEDCVEF-------AGDG----- 287

Qy        141 QWNDDRCNMKNNFICKYSDEKPAVPSRE 168
               ||| :| :|  :|||    | | |  |
Db        288 -WNDSKCELKKFWICK----KSATPCTE 310


RESULT 26
Q91ZX0
ID   Q91ZX0      PRELIMINARY;      PRT;   325 AA.
AC   Q91ZX0;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   SIGNR1 (Type II transmembrane receptor OtB7).
GN   CD209B OR SIGNR1.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6;
RX   PubMed=11581173;
RA   Park C.G., Takahara K., Umemoto E., Yashima Y., Matsubara K.,
RA   Matsuda Y., Clausen B.E., Inaba K., Steinman R.M.;
RT   "Five mouse homologues of the human dendritic cell C-type lectin, DC-
RT   SIGN.";
RL   Int. Immunol. 13:1283-1290(2001).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BALB/c; TISSUE=Thoracic lymph node;
RX   MEDLINE=22238330; PubMed=12351402;
RA   Geijtenbeek T.B.H., Groot P.C., Nolte M.A., Van Vliet S.J.,
RA   Gangaram-Panday S.T., Van Duijnhoven G.C.F., Kraal G.,
RA   Van Oosterhout A.J.M., Van Kooyk Y.;
RT   "Marginal zone macrophages express a murine homologue of DC-SIGN that
RT   captures blood-borne antigens in vivo.";
RL   Blood 100:2908-2916(2002).
DR   EMBL; AF373409; AAL13235.1; -.
DR   EMBL; AF422108; AAN31450.1; -.
DR   MGD; MGI:1916415; Cd209b.
DR   GO; GO:0016021; C:integral to membrane; IEA.
DR   GO; GO:0004872; F:receptor activity; IEA.
DR   GO; GO:0005529; F:sugar binding; IEA.
DR   InterPro; IPR002353; AntifreezeII.
DR   InterPro; IPR001304; Lectin_C.
DR   Pfam; PF00059; lectin_c; 1.
DR   PRINTS; PR00356; ANTIFREEZEII.
DR   SMART; SM00034; CLECT; 1.
DR   PROSITE; PS00615; C_TYPE_LECTIN_1; 1.
DR   PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
KW   Receptor; Transmembrane.
SQ   SEQUENCE   325 AA;  37111 MW;  9C9388407C247CA4 CRC64;

  Query Match             14.2%;  Score 158;  DB 11;  Length 325;
  Best Local Similarity   32.4%;  Pred. No. 6.7e-07;
  Matches   48;  Conservative   21;  Mismatches   45;  Indels   34;  Gaps    9;

Qy         25 YFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFWIGLRRREE 84
              ||    |:| |: :|  ||:    ||| | |::||     |::    : |  |:||   ::
Db        207 YFFSKSQR-NWNDAVTACKEVKAQLVIINSDEEQ----TFLQQTSKAKGPTWMGLSDLKK 261

Qy         85 KQSNSTACQDLYAWTDGSI--SQFRN-WYVDEP-SCGSEVCVVMYHQPSAPAGIGGPYMF 140
              : :        : | |||   |:|:  |   || : | | ||         || |
Db        262 EAT--------WLWVDGSTLSSRFQKYWNRGEPNNIGEEDCVEF-------AGDG----- 301

Qy        141 QWNDDRCNMKNNFICKYSDEKPAVPSRE 168
               ||| :| :|  :|||    | | |  |
Db        302 -WNDSKCELKKFWICK----KSATPCTE 324


RESULT 27
Q90WI7
ID   Q90WI7      PRELIMINARY;      PRT;   158 AA.
AC   Q90WI7;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   C-type lectin-like protein 2.
OS   Bungarus fasciatus (Banded krait).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
OC   Elapidae; Bungarinae; Bungarus.
OX   NCBI_TaxID=8613;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Venom gland;
RA   Zha H.-G., Zhang Y.;
RT   "cDNA cloning and characterization of C-type lectin-like proteins from
RT   Elapidae snakes.";
RL   Submitted (FEB-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF354271; AAK43585.1; -.
DR   GO; GO:0005529; F:sugar binding; IEA.
DR   GO; GO:0007157; P:heterophilic cell adhesion; IEA.
DR   InterPro; IPR002353; AntifreezeII.
DR   InterPro; IPR001304; Lectin_C.
DR   InterPro; IPR003990; Pancreatis_ac.
DR   Pfam; PF00059; lectin_c; 1.
DR   PRINTS; PR00356; ANTIFREEZEII.
DR   PRINTS; PR01504; PNCREATITSAP.
DR   SMART; SM00034; CLECT; 1.
DR   PROSITE; PS00615; C_TYPE_LECTIN_1; 1.
DR   PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
KW   Lectin.
SQ   SEQUENCE   158 AA;  18254 MW;  5F0218970DA17453 CRC64;

  Query Match             14.1%;  Score 157;  DB 13;  Length 158;
  Best Local Similarity   29.1%;  Pred. No. 3.5e-07;
  Matches   41;  Conservative   26;  Mismatches   48;  Indels   26;  Gaps    6;

Qy         20 CYKVIYFHDTSRRLNFEEAKEACR--RDGGQLVSIESEDEQKLIEKFIENLLPSDGDFWI 77
              ||||       |   :: :|:  ||  : |  | ||  : :   : ::: : |  ||: ||
Db         37 CYKVF-----SNPKSWLDAEMFCRKFKPGCHLASIHRDADSADLAEYVSDYLKDDGNVWI 91

Qy         78 GLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEP--SCGSEVCVVMYHQPSAPAGIG 135
              ||   ::|::        : |:| | | : :|   ||  |   | || ::    || |
Db         92 GLNDPQKKRT--------WVWSDRSSSNYFSWNQGEPNNSKNKEYCVHLW----APTG-- 137
```

```
Qy        136 GPYMFQWNDDRCNMKNNFICK 156
                  :|||  |   : |||:
Db        138 ---YLKWNDAPCETLHPFICQ 155


RESULT 28
Q8CJ91
ID   Q8CJ91      PRELIMINARY;      PRT;   323 AA.
AC   Q8CJ91;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   SIGNR1 alpha (Fragment).
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Swiss Webster; TISSUE=Skin;
RX   MEDLINE=22133304; PubMed=12137941;
RA   Parent S.A., Zhang T., Chrebet G., Clemas J.A., Figueroa D.J., Ky B.,
RA   Blevins R.A., Austin C.P., Rosen H.;
RT   "Molecular characterization of the murine SIGNR1 gene encoding a C-
RT   type lectin homologous to human DC-SIGN and DC-SIGNR.";
RL   Gene 293:33-46(2002).
DR   EMBL; AF424793; AAN75588.1; -.
DR   GO; GO:0005529; F:sugar binding; IEA.
DR   InterPro; IPR002353; AntifreezeII.
DR   InterPro; IPR001304; Lectin_C.
DR   Pfam; PF00059; lectin_c; 1.
DR   PRINTS; PR00356; ANTIFREEZEII.
DR   SMART; SM00034; CLECT; 1.
DR   PROSITE; PS00615; C_TYPE_LECTIN_1; 1.
DR   PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
FT   NON_TER      323     323
SQ   SEQUENCE   323 AA;  36907 MW;  3D7C512B49386C9D CRC64;

  Query Match             14.0%;  Score 156;  DB 11;  Length 323;
  Best Local Similarity   32.4%;  Pred. No. 1e-06;
  Matches   47;  Conservative   21;  Mismatches   43;  Indels   34;  Gaps    9;

Qy         25 YFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFWIGLRRREE 84
              ||   |:| |: :|  ||:    |||  | |::||     |::    : |  |:||   ::
Db        207 YFFSKSQR-NWNDAVTACKEAKAQLVIINSDEEQ----TFLQQTSKAKGPTWMGLSDLKK 261

Qy         85 KQSNSTACQDLYAWTDGSI--SQFRN-WYVDEP-SCGSEVCVVMYHQPSAPAGIGGPYMF 140
              : :        : | |||   |:|:  |   || : | | ||         || |
Db        262 EAT--------WLWVDGSTLSSRFQKYWNRGEPNNIGEEDCVEF-------AGDG----- 301

Qy        141 QWNDDRCNMKNNFICKYSDEKPAVP 165
               ||| :| :|  :|||    | | |
Db        302 -WNDSKCELKKFWICK----KSATP 321


RESULT 29
Q95244
ID   Q95244      PRELIMINARY;      PRT;   339 AA.
AC   Q95244;
DT   01-FEB-1997 (TrEMBLrel. 02, Created)
DT   01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Cartilage aggregating proteoglycan (Fragment).
OS   Sus scrofa (Pig).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX   NCBI_TaxID=9823;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=96190740; PubMed=8611178;
RA   Dudhia J., Davidson C.M., Wells T.M., Vynios D.H., Hardingham T.E.,
RA   Bayliss M.T.;
RT   "Age-related changes in the content of the C-terminal region of
RT   aggrecan in human articular cartilage.";
RL   Biochem. J. 313:933-940(1996).
DR   EMBL; X60107; CAA42701.1; -.
DR   HSSP; P23806; 1IXX.
DR   GO; GO:0005529; F:sugar binding; IEA.
DR   InterPro; IPR001304; Lectin_C.
DR   InterPro; IPR000436; Sushi_SCR_CCP.
DR   Pfam; PF00059; lectin_c; 1.
DR   Pfam; PF00084; sushi; 1.
DR   SMART; SM00032; CCP; 1.
DR   SMART; SM00034; CLECT; 1.
DR   PROSITE; PS00615; C_TYPE_LECTIN_1; 1.
DR   PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
FT   NON_TER       1       1
SQ   SEQUENCE   339 AA;  37813 MW;  D99C4B81E30A9A6F CRC64;

  Query Match             14.0%;  Score 156;  DB 6;  Length 339;
  Best Local Similarity   24.9%;  Pred. No. 1.1e-06;
  Matches   44;  Conservative   30;  Mismatches   63;  Indels   40;  Gaps    8;

Qy          9 QPVCRGGT--QRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIE 66
              | :|  |:  |  ||:  :| |   |  : :|:  ||:    | || : :||: :    :
Db        127 QKLCENGSQIQGHCYR--HFPD---RATWVDAESQCRKQQSHLSSIVTPEEQEFVNNNAQ 181

Qy         67 NLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEP----SCGSEVCV 122
                  ||||         |    :  : |:||   || :|  ::|    : | :  |
Db        182 EYQ------WIGL--------NDKTIEGDFRWSDGHSLQFESWSPNQPDNFFATGEDCVV 227

Qy        123 VMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYS----DEKPAVPSREAEGEETE 175
              :::|:              :|||  || :  | ||       | | |       |:: :
Db        228 MIWHEKG----------EWNDVPCNYQLPFTCKKGTVACGEPPVVEHARIFGKKKD 273


RESULT 30
Q8BGZ0
ID   Q8BGZ0      PRELIMINARY;      PRT;   293 AA.
AC   Q8BGZ0;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   SIGNR1 beta (Fragment).
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Swiss Webster; TISSUE=Spleen;
RX   MEDLINE=22133304; PubMed=12137941;
RA   Parent S.A., Zhang T., Chrebet G., Clemas J.A., Figueroa D.J., Ky B.,
RA   Blevins R.A., Austin C.P., Rosen H.;
RT   "Molecular characterization of the murine SIGNR1 gene encoding a C-
RT   type lectin homologous to human DC-SIGN and DC-SIGNR.";
RL   Gene 293:33-46(2002).
DR   EMBL; AF424797; AAN75592.1; -.
DR   EMBL; AF424798; AAN75593.1; -.
DR   GO; GO:0005529; F:sugar binding; IEA.
DR   InterPro; IPR002353; AntifreezeII.
DR   InterPro; IPR001304; Lectin_C.
DR   Pfam; PF00059; lectin_c; 1.
DR   PRINTS; PR00356; ANTIFREEZEII.
DR   SMART; SM00034; CLECT; 1.
DR   PROSITE; PS00615; C_TYPE_LECTIN_1; 1.
DR   PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
FT   NON_TER      293     293
SQ   SEQUENCE   293 AA;  33698 MW;  6113551A91D09BCA CRC64;

  Query Match             13.9%;  Score 155;  DB 11;  Length 293;
  Best Local Similarity   32.4%;  Pred. No. 1.1e-06;
  Matches   47;  Conservative   21;  Mismatches   43;  Indels   34;  Gaps    9;

Qy         25 YFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFWIGLRRREE 84
```

```
                ||   |:| |: :|  ||:    |||  | |::||     |::    : |  |:||   ::
Db        177 YFFSKSQR-NWNDAVTACKEVKAQLVIINSDEEQ----TFLQQTSKAKGPTWMGLSDLKK 231

Qy         85 KQSNSTACQDLYAWTDGSI--SQFRN-WYVDEP-SCGSEVCVVMYHQPSAPAGIGGPYMF 140
                : :         : | |||   |:|:  |   || : | | ||          || |
Db        232 EAT--------WLWVDGSTLSSRFQKYWNRGEPNNIGEEDCVEF-------AGDG----- 271

Qy        141 QWNDDRCNMKNNFICKYSDEKPAVP 165
                ||| :| :|  :|||    | | |
Db        272 -WNDSKCELKKFWICK----KSATP 291


RESULT 31
Q8CJ94
ID   Q8CJ94      PRELIMINARY;      PRT;   323 AA.
AC   Q8CJ94;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   SIGNR1 (Fragment).
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Swiss Webster; TISSUE=Skin;
RX   MEDLINE=22133304; PubMed=12137941;
RA   Parent S.A., Zhang T., Chrebet G., Clemas J.A., Figueroa D.J., Ky B.,
RA   Blevins R.A., Austin C.P., Rosen H.;
RT   "Molecular characterization of the murine SIGNR1 gene encoding a C-
RT   type lectin homologous to human DC-SIGN and DC-SIGNR.";
RL   Gene 293:33-46(2002).
DR   EMBL; AF424790; AAN75585.1; -.
DR   GO; GO:0005529; F:sugar binding; IEA.
DR   InterPro; IPR002353; AntifreezeII.
DR   InterPro; IPR001304; Lectin_C.
DR   Pfam; PF00059; lectin_c; 1.
DR   PRINTS; PR00356; ANTIFREEZEII.
DR   SMART; SM00034; CLECT; 1.
DR   PROSITE; PS00615; C_TYPE_LECTIN_1; 1.
DR   PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
FT   NON_TER     323     323
SQ   SEQUENCE   323 AA;  36819 MW;  EDA0E46E1969AE54 CRC64;

  Query Match             13.9%;  Score 155;  DB 11;  Length 323;
  Best Local Similarity   32.4%;  Pred. No. 1.3e-06;
  Matches   47;  Conservative   21;  Mismatches   43;  Indels   34;  Gaps    9;

Qy         25 YFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFWIGLRRREE 84
                ||   |:| |: :|  ||:    |||  | |::||     |::    : |  |:||   ::
Db        207 YFFSKSQR-NWNDAVTACKEVKAQLVIINSDEEQ----TFLQQTSKAKGPTWMGLSDLKK 261

Qy         85 KQSNSTACQDLYAWTDGSI--SQFRN-WYVDEP-SCGSEVCVVMYHQPSAPAGIGGPYMF 140
                : :         : | |||   |:|:  |   || : | | ||          || |
Db        262 EAT--------WLWVDGSTLSSRFQKYWNRGEPNNIGEEDCVEF-------AGDG----- 301

Qy        141 QWNDDRCNMKNNFICKYSDEKPAVP 165
                ||| :| :|  :|||    | | |
Db        302 -WNDSKCGLKKFWICK----KSATP 321


RESULT 32
Q8CJ93
ID   Q8CJ93      PRELIMINARY;      PRT;   323 AA.
AC   Q8CJ93;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   SIGNR1 alpha (Fragment).
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Swiss Webster; TISSUE=Spleen;
RX   MEDLINE=22133304; PubMed=12137941;
RA   Parent S.A., Zhang T., Chrebet G., Clemas J.A., Figueroa D.J., Ky B.,
RA   Blevins R.A., Austin C.P., Rosen H.;
RT   "Molecular characterization of the murine SIGNR1 gene encoding a C-
RT   type lectin homologous to human DC-SIGN and DC-SIGNR.";
RL   Gene 293:33-46(2002).
DR   EMBL; AF424791; AAN75586.1; -.
DR   GO; GO:0005529; F:sugar binding; IEA.
DR   InterPro; IPR002353; AntifreezeII.
DR   InterPro; IPR001304; Lectin_C.
DR   Pfam; PF00059; lectin_c; 1.
DR   PRINTS; PR00356; ANTIFREEZEII.
DR   SMART; SM00034; CLECT; 1.
DR   PROSITE; PS00615; C_TYPE_LECTIN_1; 1.
DR   PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
FT   NON_TER     323     323
SQ   SEQUENCE   323 AA;  36908 MW;  5289B6382D9405B6 CRC64;

  Query Match             13.9%;  Score 155;  DB 11;  Length 323;
  Best Local Similarity   32.4%;  Pred. No. 1.3e-06;
  Matches   47;  Conservative   21;  Mismatches   43;  Indels   34;  Gaps    9;

Qy         25 YFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFWIGLRRREE 84
                ||   |:| |: :|  ||:    |||  | |::||     |::    : |  |:||   ::
Db        207 YFFSKSQR-NWNDAVTACKEVKAQLVIINSDEEQ----TFLQQTSKAKGPTWMGLSDLKK 261

Qy         85 KQSNSTACQDLYAWTDGSI--SQFRN-WYVDEP-SCGSEVCVVMYHQPSAPAGIGGPYMF 140
                : :         : | |||   |:|:  |   || : | | ||          || |
Db        262 EAT--------WLWVDGSTLSSRFQKYWNRGEPNNIGEEDCVEF-------AGDG----- 301

Qy        141 QWNDDRCNMKNNFICKYSDEKPAVP 165
                ||| :| :|  :|||    | | |
Db        302 -WNDSKCELKKFWICK----KSATP 321


RESULT 33
Q8CJ88
ID   Q8CJ88      PRELIMINARY;      PRT;   323 AA.
AC   Q8CJ88;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   SIGNR1 alpha (Fragment).
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Swiss Webster; TISSUE=Spleen;
RX   MEDLINE=22133304; PubMed=12137941;
RA   Parent S.A., Zhang T., Chrebet G., Clemas J.A., Figueroa D.J., Ky B.,
RA   Blevins R.A., Austin C.P., Rosen H.;
RT   "Molecular characterization of the murine SIGNR1 gene encoding a C-
RT   type lectin homologous to human DC-SIGN and DC-SIGNR.";
RL   Gene 293:33-46(2002).
DR   EMBL; AF424796; AAN75591.1; -.
DR   GO; GO:0005529; F:sugar binding; IEA.
DR   InterPro; IPR002353; AntifreezeII.
DR   InterPro; IPR001304; Lectin_C.
DR   Pfam; PF00059; lectin_c; 1.
DR   PRINTS; PR00356; ANTIFREEZEII.
DR   SMART; SM00034; CLECT; 1.
DR   PROSITE; PS00615; C_TYPE_LECTIN_1; 1.
DR   PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
FT   NON_TER     323     323
SQ   SEQUENCE   323 AA;  36921 MW;  2E807C247CA4B163 CRC64;

  Query Match             13.9%;  Score 155;  DB 11;  Length 323;
```

```
  Best Local Similarity   32.4%;  Pred. No. 1.3e-06;
  Matches   47;  Conservative   21;  Mismatches   43;  Indels   34;  Gaps    9;

Qy         25 YFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFWIGLRRREE 84
              ||   |:| |: :|  ||:    ||| | |::||     |::    : |  |:||   ::
Db        207 YFFSKSQR-NWNDAVTACKEVKAQLVIINSDEEQ----TFLQQTSKAKGPTWMGLSDLKK 261

Qy         85 KQSNSTACQDLYAWTDGSI--SQFRN-WYVDEP-SCGSEVCVVMYHQPSAPAGIGGPYMF 140
              : :        : | |||   |:|:  |   || : | | ||         || |
Db        262 EAT--------WLWVDGSTLSSRFQKYWNRGEPNNIGEEDCVEF-------AGDG----- 301

Qy        141 QWNDDRCNMKNNFICKYSDEKPAVP 165
               ||| :| :|  :|||    | | |
Db        302 -WNDSKCELKKFWICK----KSATP 321


RESULT 34
Q7TMA7
ID   Q7TMA7      PRELIMINARY;      PRT;   379 AA.
AC   Q7TMA7;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Mannose receptor-like isoform 3 (Mannose receptor-like isoform
DE   5).
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Cerebellum;
RA   Jang J.S. Sr., Ackerman S.L.;
RT   "Alternatively spliced isoform 3 of mannose receptor precursor-like
RT   gene.";
RL   Submitted (JAN-2003) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Cerebellum;
RA   Jang J.S. Sr., Ackerman S.L.;
RT   "Alternatively spliced isoform 5 of mannose receptor precursor-like
RT   gene.";
RL   Submitted (JAN-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY223870; AAP22984.1; -.
DR   EMBL; AY223872; AAP22986.1; -.
KW   Receptor.
SQ   SEQUENCE   379 AA;  42164 MW;  98D94076795F85D5 CRC64;

  Query Match             13.9%;  Score 155;  DB 11;  Length 379;
  Best Local Similarity   27.8%;  Pred. No. 1.6e-06;
  Matches   47;  Conservative   23;  Mismatches   59;  Indels   40;  Gaps    7;

Qy         14 GGTQRPCYKVIYFHDTSRRLN-----------FEEAKEACRRDGGQLVSIESEDEQKLI 61
              || |         ||    |::           :  ||  |:  || |  |||:  | ::
Db        220 GGAQCATKVGFPFHTCDLRIDGDCFMVSPEADTYYGAKMKCQGKGGVLAQIESQKVQDIL 279

Qy         62 EKFIENLLP----SDGD-----FWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVD 112
                ::  |      :| |     |||||  :        | :| : || |     | ::
Db        280 AFYLGRLETTNEVTDSDFETKNFWIGLTYK--------AAKDSFRWTTGEHQSFTSFAFG 331

Qy        113 EP-SCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDE 160
              :| : |   || |  | ||         | ||| ||  :| :||::: :
Db        332 QPDNQGFGNCVEM--QASA--------AFNWNDQRCKTRNRYICQFAQK 370


RESULT 35
Q7TSP9
ID   Q7TSP9      PRELIMINARY;      PRT;   473 AA.
AC   Q7TSP9;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Mannose receptor-like isoform 6.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Cerebellum;
RA   Jang J.S., Ackerman S.L.;
RT   "cDNA sequence of a mannose receptor precursor-like gene.";
RL   Submitted (JAN-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY223868; AAP20650.1; -.
KW   Receptor.
SQ   SEQUENCE   473 AA;  52563 MW;  707C43A6B3C2E206 CRC64;

  Query Match             13.9%;  Score 155;  DB 11;  Length 473;
  Best Local Similarity   27.8%;  Pred. No. 2e-06;
  Matches   47;  Conservative   23;  Mismatches   59;  Indels   40;  Gaps    7;

Qy         14 GGTQRPCYKVIYFHDTSRRLN-----------FEEAKEACRRDGGQLVSIESEDEQKLI 61
              || |         ||    |::           :  ||  |:  || |  |||:  | ::
Db        314 GGAQCATKVGFPFHTCDLRIDGDCFMVSPEADTYYGAKMKCQGKGGVLAQIESQKVQDIL 373

Qy         62 EKFIENLLP----SDGD-----FWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVD 112
                ::  |      :| |     |||||  :        | :| : || |     | ::
Db        374 AFYLGRLETTNEVTDSDFETKNFWIGLTYK--------AAKDSFRWTTGEHQSFTSFAFG 425

Qy        113 EP-SCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDE 160
              :| : |   || |  | ||         | ||| ||  :| :||::: :
Db        426 QPDNQGFGNCVEM--QASA--------AFNWNDQRCKTRNRYICQFAQK 464


RESULT 36
Q7TSQ7
ID   Q7TSQ7      PRELIMINARY;      PRT;   477 AA.
AC   Q7TSQ7;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Mannose receptor-like.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Cerebellum;
RA   Jang J.S., Ackerman S.L.;
RT   "cDNA sequence of a mannose receptor precursor-like gene.";
RL   Submitted (JAN-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY223868; AAP20650.1; -.
KW   Receptor.
SQ   SEQUENCE   477 AA;  52490 MW;  E24C5071B4137148 CRC64;

  Query Match             13.9%;  Score 155;  DB 11;  Length 477;
  Best Local Similarity   27.8%;  Pred. No. 2.1e-06;
  Matches   47;  Conservative   23;  Mismatches   59;  Indels   40;  Gaps    7;

Qy         14 GGTQRPCYKVIYFHDTSRRLN-----------FEEAKEACRRDGGQLVSIESEDEQKLI 61
              || |         ||    |::           :  ||  |:  || |  |||:  | ::
Db        318 GGAQCATKVGFPFHTCDLRIDGDCFMVSPEADTYYGAKMKCQGKGGVLAQIESQKVQDIL 377

Qy         62 EKFIENLLP----SDGD-----FWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVD 112
                ::  |      :| |     |||||  :        | :| : || |     | ::
Db        378 AFYLGRLETTNEVTDSDFETKNFWIGLTYK--------AAKDSFRWTTGEHQSFTSFAFG 429

Qy        113 EP-SCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDE 160
              :| : |   || |  | ||         | ||| ||  :| :||::: :
Db        430 QPDNQGFGNCVEM--QASA--------AFNWNDQRCKTRNRYICQFAQK 468


RESULT 37
Q7TSQ0
```

```
ID   Q7TSQ0      PRELIMINARY;      PRT;   504 AA.
AC   Q7TSQ0;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Mannose receptor-like isoform 4.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Cerebellum;
RA   Jang J.S. Sr., Ackerman S.L.;
RT   "Alternatively spliced isoform 4 of mannose receptor precursor-like
RT   gene.";
RL   Submitted (JAN-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY223871; AAP22985.1; -.
KW   Receptor.
SQ   SEQUENCE   504 AA;  55995 MW;  A964C0C319D77B8C CRC64;

  Query Match             13.9%;  Score 155;  DB 11;  Length 504;
  Best Local Similarity   27.8%;  Pred. No. 2.2e-06;
  Matches   47;  Conservative   23;  Mismatches   59;  Indels   40;  Gaps    7;

Qy         14 GGTQRPCYKVIYFHDTSRRLN------------FEEAKEACRRDGGQLVSIESEDEQKLI 61
              || |         ||     |::            :  ||  |:  || |  |||:  | ::
Db        345 GGAQCATKVGFPFHTCDLRIDGDCFMVSPEADTYYGAKMKCQGKGGVLAQIESQKVQDIL 404

Qy         62 EKFIENLLP----SDGD-----FWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVD 112
                ::  |       :| |      |||||  :         | :| : || |     | ::
Db        405 AFYLGRLETTNEVTDSDFETKNFWIGLTYK--------AAKDSFRWTTGEHQSFTSFAFG 456

Qy        113 EP-SCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDE 160
              :| : |   || |  | ||         | ||| ||  :| :|||::: :
Db        457 QPDNQGFGNCVEM--QASA--------AFNWNDQRCKTRNRYICQFAQK 495


RESULT 38
Q7TSQ1
ID   Q7TSQ1      PRELIMINARY;      PRT;   534 AA.
AC   Q7TSQ1;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Mannose receptor-like isoform 2.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Cerebellum;
RA   Jang J.S. Sr., Ackerman S.L.;
RT   "Alternatively spliced isoform 2 of mannose receptor precursor-like.";
RL   Submitted (JAN-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY223869; AAP22983.1; -.
KW   Receptor.
SQ   SEQUENCE   534 AA;  58585 MW;  9B0AF013370793FF CRC64;

  Query Match             13.9%;  Score 155;  DB 11;  Length 534;
  Best Local Similarity   27.8%;  Pred. No. 2.4e-06;
  Matches   47;  Conservative   23;  Mismatches   59;  Indels   40;  Gaps    7;

Qy         14 GGTQRPCYKVIYFHDTSRRLN------------FEEAKEACRRDGGQLVSIESEDEQKLI 61
              || |         ||     |::            :  ||  |:  || |  |||:  | ::
Db        375 GGAQCATKVGFPFHTCDLRIDGDCFMVSPEADTYYGAKMKCQGKGGVLAQIESQKVQDIL 434

Qy         62 EKFIENLLP----SDGD-----FWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVD 112
                ::  |       :| |      |||||  :         | :| : || |     | ::
Db        435 AFYLGRLETTNEVTDSDFETKNFWIGLTYK--------AAKDSFRWTTGEHQSFTSFAFG 486

Qy        113 EP-SCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDE 160
              :| : |   || |  | ||         | ||| ||  :| :|||::: :
Db        487 QPDNQGFGNCVEM--QASA--------AFNWNDQRCKTRNRYICQFAQK 525


RESULT 39
Q90WM2
ID   Q90WM2      PRELIMINARY;      PRT;  1152 AA.
AC   Q90WM2;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Brevican soluble core protein.
OS   Xenopus laevis (African clawed frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
OC   Xenopodinae; Xenopus.
OX   NCBI_TaxID=8355;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21184128; PubMed=11287204;
RA   Sander V., Mullegger J., Lepperdinger G.;
RT   "Xenopus brevican is expressed in the notochord and the brain during
RT   early embryogenesis.";
RL   Mech. Dev. 102:251-253(2001).
DR   EMBL; AF325324; AAK40085.1; -.
DR   GO; GO:0005540; F:hyaluronic acid binding; IEA.
DR   GO; GO:0005529; F:sugar binding; IEA.
DR   GO; GO:0007155; P:cell adhesion; IEA.
DR   InterPro; IPR002353; AntifreezeII.
DR   InterPro; IPR000742; EGF_2.
DR   InterPro; IPR006209; EGF_like.
DR   InterPro; IPR006210; IEGF.
DR   InterPro; IPR003599; Ig.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR001304; Lectin_C.
DR   InterPro; IPR000538; Link.
DR   InterPro; IPR000436; Sushi_SCR_CCP.
DR   Pfam; PF00008; EGF; 1.
DR   Pfam; PF00059; lectin_c; 1.
DR   Pfam; PF00084; sushi; 1.
DR   Pfam; PF00193; Xlink; 2.
DR   PRINTS; PR00356; ANTIFREEZEII.
DR   PRINTS; PR01265; LINKMODULE.
DR   ProDom; PD000918; Link; 2.
DR   SMART; SM00032; CCP; 1.
DR   SMART; SM00034; CLECT; 1.
DR   SMART; SM00181; EGF; 1.
DR   SMART; SM00409; IG; 1.
DR   SMART; SM00445; LINK; 2.
DR   PROSITE; PS00615; C_TYPE_LECTIN_1; 1.
DR   PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
DR   PROSITE; PS00022; EGF_1; 1.
DR   PROSITE; PS01186; EGF_2; 1.
DR   PROSITE; PS50835; IG_LIKE; 1.
DR   PROSITE; PS00290; IG_MHC; 1.
DR   PROSITE; PS01241; LINK; 2.
KW   EGF-like domain.
SQ   SEQUENCE   1152 AA;  126860 MW;  196947BD8F9B216D CRC64;

  Query Match             13.9%;  Score 155;  DB 13;  Length 1152;
  Best Local Similarity   30.6%;  Pred. No. 5.9e-06;
  Matches   49;  Conservative   17;  Mismatches   56;  Indels   38;  Gaps    9;

Qy         20 CYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDF-WIG 78
              |||  :||    | ::|||:  ||  || | || : :||     |: |      |: | |
Db        973 CYK--HFH---ARRSWEEAENFCREAGGHLTSIMTPEEQ----AFLSN---KYNDYQWTG 1020

Qy         79 LRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEPS---CGSEVCVVMYHQPSAPAGIG 135
              |       |    :  : |:||:   | ||   :|       | ||||
Db        1021 L--------NDRTIEGDFQWSDGNPLLFENWAHGQPDSYFLSGENCVVMV---------- 1062

Qy        136 GPYMFQWNDDRCNMKNNFICKYS----DEKPAVPSREAEG 171
```

```
                |    :|:|  ||    |:||         | ||:    |
Db       1063 GHNEGKWSDVPCNYHLPFVCKMGLVSCGPPPDVPNATMYG 1102
```

RESULT 40
Q8CJ89
ID   Q8CJ89   PRELIMINARY;   PRT;   322 AA.
AC   Q8CJ89;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   SIGNR1 alpha (Fragment).
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Swiss Webster; TISSUE=Skin;
RX   MEDLINE=22133304; PubMed=12137941;
RA   Parent S.A., Zhang T., Chrebet G., Clemas J.A., Figueroa D.J., Ky B.,
RA   Blevins R.A., Austin C.P., Rosen H.;
RT   "Molecular characterization of the murine SIGNR1 gene encoding a C-
RT   type lectin homologous to human DC-SIGN and DC-SIGNR.";
RL   Gene 293:33-46(2002).
DR   EMBL; AF424795; AAN75590.1; -.
DR   GO; GO:0005529; F:sugar binding; IEA.
DR   InterPro; IPR002353; AntifreezeII.
DR   InterPro; IPR001304; Lectin_C.
DR   Pfam; PF00059; lectin_c; 1.
DR   PRINTS; PR00356; ANTIFREEZEII.
DR   SMART; SM00034; CLECT; 1.
DR   PROSITE; PS00615; C_TYPE_LECTIN_1; 1.
DR   PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
FT   NON_TER   322   322
SQ   SEQUENCE   322 AA; 36832 MW; E30AA47CA4B16362 CRC64;

  Query Match   13.8%; Score 154; DB 11; Length 322;
  Best Local Similarity   32.6%; Pred. No. 1.6e-06;
  Matches   45; Conservative   21; Mismatches   42; Indels   30; Gaps   8;

```
Qy          25 YFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFWIGLRRREE 84
               ||   |:| |: :|  ||:    |||  | |::||     |::    : |  |:||   ::
Db         207 YFFSKSQR-NWNDAVTACKEVKAQLVIINSDEEQ----TFLQQTSKAKGPTWMGLSDLKK 261

Qy          85 KQSNSTACQDLYAWTDGSI--SQFRN-WYVDEP-SCGSEVCVVMYHQPSAPAGIGGPYMF 140
               : :        : | |||    |:|:  |    || : | | ||          || |
Db         262 EAT--------WLWVDGSTLSSRFQKYWNRGEPNNIGEEDCVEF-------AGDG----- 301

Qy         141 QWNDDRCNMKNNFICKYS 158
                |||  :| :|  :||| |
Db         302 -WNDSKCELKKFWICKKS 318
```

RESULT 41
Q8CJ92
ID   Q8CJ92   PRELIMINARY;   PRT;   323 AA.
AC   Q8CJ92;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   SIGNR1 alpha (Fragment).
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Swiss Webster; TISSUE=Spleen;
RX   MEDLINE=22133304; PubMed=12137941;
RA   Parent S.A., Zhang T., Chrebet G., Clemas J.A., Figueroa D.J., Ky B.,
RA   Blevins R.A., Austin C.P., Rosen H.;
RT   "Molecular characterization of the murine SIGNR1 gene encoding a C-
RT   type lectin homologous to human DC-SIGN and DC-SIGNR.";
RL   Gene 293:33-46(2002).
DR   EMBL; AF424792; AAN75587.1; -.
DR   GO; GO:0005529; F:sugar binding; IEA.
DR   InterPro; IPR002353; AntifreezeII.
DR   InterPro; IPR001304; Lectin_C.
DR   Pfam; PF00059; lectin_c; 1.
DR   PRINTS; PR00356; ANTIFREEZEII.
DR   SMART; SM00034; CLECT; 1.
DR   PROSITE; PS00615; C_TYPE_LECTIN_1; 1.
DR   PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
FT   NON_TER   323   323
SQ   SEQUENCE   323 AA; 36879 MW; 2B2A9979E5AD34AE CRC64;

  Query Match   13.8%; Score 154; DB 11; Length 323;
  Best Local Similarity   31.7%; Pred. No. 1.6e-06;
  Matches   46; Conservative   22; Mismatches   43; Indels   34; Gaps   9;

```
Qy          25 YFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFWIGLRRREE 84
               ||   |:| |: :|  ||:    |||:|  ::||     |::    : |  |:||   ::
Db         207 YFFSKSQR-NWNDAVTACKEVKAQLVTINGDEEQ----TFLQQTSKAKGPTWMGLSDLKK 261

Qy          85 KQSNSTACQDLYAWTDGSI--SQFRN-WYVDEP-SCGSEVCVVMYHQPSAPAGIGGPYMF 140
               : :        : | |||    |:|:  |    || : | | ||          || |
Db         262 EAT--------WLWVDGSTLSSRFQKYWNRGEPNNIGEEDCVEF-------AGDG----- 301

Qy         141 QWNDDRCNMKNNFICKYSDEKPAVP 165
                ||| :| :|  :|||    | | |
Db         302 -WNDSKCELKKFWICK----KSATP 321
```

RESULT 42
Q7Z5K9
ID   Q7Z5K9   PRELIMINARY;   PRT;   446 AA.
AC   Q7Z5K9;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Unnamed secretory protein.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Ding P., Han W., Rui M., Wang Y., Zhang Y., Song Q., Ma D.;
RL   Submitted (JUN-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF521893; AAP80866.1; -.
SQ   SEQUENCE   446 AA; 49574 MW; EF34732BF9ECD0C1 CRC64;

  Query Match   13.7%; Score 153; DB 4; Length 446;
  Best Local Similarity   26.9%; Pred. No. 2.9e-06;
  Matches   46; Conservative   27; Mismatches   54; Indels   44; Gaps   9;

```
Qy          14 GGTQRPCYKVIY--FHDTSRRLN-----------FEEAKEACRRDGGQLVSIESEDEQK 59
               || |  |   ::  ||     |::           :  |:  |:| || |  |:|:  |
Db         287 GGAQ--CATKVHFPFHTCDLRIDGDCFMVSSEADTYYRARMKCQRKGGVLAQIKSQKVQD 344

Qy          60 LIEKFIENL------LPSD---GDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWY 110
               ::  ::  |      : ||    :|||||  :  |       | : |  |     | ::
Db         345 ILAFYLGRLETTNEVIDSDFETRNFWIGLTYKTAK--------DSFRWATGEHQAFTSFA 396

Qy         111 VDEP-SCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDE 160
                 :| : |   || :  | ||          | ||| ||  :| :||::: |
Db         397 FGQPDNHGFGNCVEL--QASA--------AFNWNDQRCKTRNRYICQFAQE 437
```

RESULT 43
Q95LG3
ID   Q95LG3   PRELIMINARY;   PRT;   485 AA.
AC   Q95LG3;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)

```
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   E-selectin.
OS   Odocoileus hemionus (Mule deer) (Black-tailed deer).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Cervoidea;
OC   Cervidae; Odocoileinae; Odocoileus.
OX   NCBI_TaxID=9872;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21421234; PubMed=11529941;
RA   Hedges J.F., Demaula C.D., Moore B.D., McLaughlin B.E., Simon S.I.,
RA   MacLachlan N.J.;
RT   "Characterization of equine E-selectin.";
RL   Immunology 103:498-504(2001).
DR   EMBL; AF307970; AAK48710.1; -.
DR   GO; GO:0016020; C:membrane; IEA.
DR   GO; GO:0005194; F:cell adhesion molecule activity; IEA.
DR   GO; GO:0005529; F:sugar binding; IEA.
DR   GO; GO:0007155; P:cell adhesion; IEA.
DR   GO; GO:0006952; P:defense response; IEA.
DR   GO; GO:0007157; P:heterophilic cell adhesion; IEA.
DR   InterPro; IPR000742; EGF_2.
DR   InterPro; IPR006209; EGF_like.
DR   InterPro; IPR006210; IEGF.
DR   InterPro; IPR001304; Lectin_C.
DR   InterPro; IPR002396; Selectin.
DR   InterPro; IPR000436; Sushi_SCR_CCP.
DR   Pfam; PF00008; EGF; 1.
DR   Pfam; PF00059; lectin_c; 1.
DR   Pfam; PF00084; sushi; 4.
DR   PRINTS; PR00343; SELECTIN.
DR   SMART; SM00032; CCP; 4.
DR   SMART; SM00034; CLECT; 1.
DR   SMART; SM00181; EGF; 1.
DR   PROSITE; PS00615; C_TYPE_LECTIN_1; 1.
DR   PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
DR   PROSITE; PS00022; EGF_1; 1.
DR   PROSITE; PS01186; EGF_2; 1.
KW   EGF-like domain; Lectin; Selectin.
SQ   SEQUENCE   485 AA;  53247 MW;  69959199EAFE9980 CRC64;

  Query Match             13.7%;  Score 152.5;  DB 6;  Length 485;
  Best Local Similarity   26.0%;  Pred. No. 3.6e-06;
  Matches   40;  Conservative   32;  Mismatches   57;  Indels   25;  Gaps    6;

Qy         26 FHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFWIGLRRREEK 85
              :| :   : ||||:: |::    ||:|::::| | :     :   |   :|||:|:
Db         25 YHASKENMTFEEARDYCQKTYTALVAIQNQEEIKYLNSTFSH---SPSYYWIGIRK---- 77

Qy         86 QSNSTACQDLYAW--TDGSIS-QFRNWYVDEPS--CGSEVCVVMYHQPSAPAGIGGPYMF 140
                    |:: |  |: |:: :  ||   ||:     | || :| :    :|
Db         78 ------INDIWTWIGTNKSLTKEATNWAPGEPNNKQSDEDCVEIYIKREKDSG------- 124

Qy        141 QWNDDRCNMKNNFICKYSDEKPAVPSREAEGEET 174
              :|||:||  :   :|  :   |   |   ||
Db        125 KWNDERCTKRKLALCYKAACTPTPCSSHGECVET 158


RESULT 44
Q90WI6
ID   Q90WI6      PRELIMINARY;      PRT;   158 AA.
AC   Q90WI6;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   C-type lectin-like protein 1.
OS   Bungarus multicinctus (Many-banded krait).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
OC   Elapidae; Bungarinae; Bungarus.
OX   NCBI_TaxID=8616;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Venom gland;
RA   Zha H.-G., Zhang Y.;
RT   "cDNA cloning and characterization of C-type lectin-like proteins from
RT   Elapidae snakes.";
RL   Submitted (FEB-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF354272; AAK43586.1; -.
DR   GO; GO:0005529; F:sugar binding; IEA.
DR   GO; GO:0007157; P:heterophilic cell adhesion; IEA.
DR   InterPro; IPR001304; Lectin_C.
DR   InterPro; IPR003990; Pancreatis_ac.
DR   Pfam; PF00059; lectin_c; 1.
DR   PRINTS; PR01504; PNCREATITSAP.
DR   SMART; SM00034; CLECT; 1.
DR   PROSITE; PS00615; C_TYPE_LECTIN_1; 1.
DR   PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
KW   Lectin.
SQ   SEQUENCE   158 AA;  18706 MW;  66B71A29D1048805 CRC64;

  Query Match             13.6%;  Score 152;  DB 13;  Length 158;
  Best Local Similarity   28.4%;  Pred. No. 1.1e-06;
  Matches   40;  Conservative   22;  Mismatches   53;  Indels   26;  Gaps    5;

Qy         20 CYKVIYFHDTSRRLNFEEAKEACR--RDGGQLVSIESEDEQKLIEKFIENLLPSDGDFWI 77
              ||||   | |     : :|:: ||  : |  | |: |  :     ::| : |   |  ||
Db         37 CYKVFSKHKT-----WFDAEKYCRKFKPGCHLASLHSNADAVEFSEYISDYLTGRGHVWI 91

Qy         78 GLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEPS--CGSEVCVVMYHQPSAPAGIG 135
              |||  ::|          :: ||| | : |  |  ::|     :| || :         :
Db         92 GLRDTKKKY--------IWEWTDRSRTDFLPWRKNQPDHFNNNEFCVEI---------VN 134

Qy        136 GPYMFQWNDDRCNMKNNFICK 156
                   |||| |     |:|:
Db        135 FTGYLQWNDDNCAALRPFLCQ 155


RESULT 45
Q8IXK1
ID   Q8IXK1      PRELIMINARY;      PRT;   652 AA.
AC   Q8IXK1;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   C1q receptor protein precursor.
GN   CD93.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Steinberger P., Stoeckl J., Wille S., Szekeres A., Prager E.,
RA   Staffler G., Kuenig S., Kohl P., Majdic O., Knapp W., Stockinger H.;
RT   "Identification of CD93 as the C1q receptor protein (C1qRp) by
RT   retroviral expression cloning.";
RL   Submitted (OCT-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AJ295142; CAC82720.1; -.
DR   GO; GO:0016021; C:integral to membrane; IEA.
DR   GO; GO:0005509; F:calcium ion binding; IEA.
DR   GO; GO:0004872; F:receptor activity; IEA.
DR   GO; GO:0005529; F:sugar binding; IEA.
DR   GO; GO:0007596; P:blood coagulation; IEA.
DR   InterPro; IPR000152; Asx_hydroxyl_S.
DR   InterPro; IPR001881; EGF_Ca.
DR   InterPro; IPR006209; EGF_like.
DR   InterPro; IPR006210; IEGF.
DR   InterPro; IPR001304; Lectin_C.
DR   InterPro; IPR001187; Tissue_factor.
DR   Pfam; PF00008; EGF; 5.
DR   Pfam; PF00059; lectin_c; 1.
DR   Pfam; PF01108; Tissue_fac; 1.
DR   SMART; SM00034; CLECT; 1.
DR   SMART; SM00181; EGF; 5.
DR   SMART; SM00179; EGF_CA; 5.
```

```
DR   PROSITE; PS00010; ASX_HYDROXYL; 3.
DR   PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
DR   PROSITE; PS01186; EGF_2; 3.
DR   PROSITE; PS01187; EGF_CA; 3.
KW   Signal; Receptor.
FT   SIGNAL        1     21       POTENTIAL.
FT   CHAIN        22    652       C1Q RECEPTOR PROTEIN.
SQ   SEQUENCE   652 AA;  68521 MW;  97549BA62CAF225E CRC64;

  Query Match             13.6%;  Score 151.5;  DB 4;  Length 652;
  Best Local Similarity   24.4%;  Pred. No. 6.5e-06;
  Matches   50;  Conservative   38;  Mismatches   74;  Indels   43;  Gaps    9;

Qy          5 LLSGQP-----------VCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIE 53
              ||  ||           || |     ||        | :|:  ||:  | ::|| | :::
Db         13 LLLTQPGAGTGADTEAVVCVG---TACYTA-----HSGKLSAAEAQNHCNQNGGNLATVK 64

Qy         54 SEDEQKLIEKFIENLLPSD-------GDFWIGLRRREEKQSNSTACQDLYAWT-DGSISQ 105
              |::| : ::: :  ||  :         |||||:| : |  : :     ::|   |  :
Db         65 SKEEAQHVQRVLAQLLRREAALTARMSKFWIGLQREKGKCLDPSLPLKGFSWVGGGEDTP 124

Qy        106 FRNWYVD-EPSCGSEVCVVM---YHQPSAPAGIGGPYMFQWNDDRCNMKNN-------FI 154
              : ||: :   || |: || :      ||  |: :      :|::  |    :       |:
Db        125 YSNWHKELRNSCISKRCVSLLLDLSQPLLPSRLP-----KWSEGPCGSPGSPGSNIEGFV 179

Qy        155 CKYSDEKPAVPSREAEGEETELTTP 179
              ||:| :    |       :  |||
Db        180 CKFSFKGMCRPLALGGPGQVTYTTP 204


RESULT 46
Q8HY12
ID   Q8HY12      PRELIMINARY;      PRT;   399 AA.
AC   Q8HY12;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Putative CD209L1 protein.
GN   CD209L1.
OS   Hylobates lar (Common gibbon).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hylobatidae; Hylobates.
OX   NCBI_TaxID=9580;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=B23;
RX   PubMed=12477827;
RA   Bashirova A.A., Wu L., Cheng J., Martin T.D., Martin M.P.,
RA   Benveniste R.E., Lifson J.D., KewalRamani V.N., Hughes A.,
RA   Carrington M.;
RT   "Novel Member of the CD209 (DC-SIGN) Gene Family in Primates.";
RL   J. Virol. 77:217-227(2003).
DR   EMBL; AY078813; AAL89528.1; -.
DR   EMBL; AY078807; AAL89528.1; JOINED.
DR   EMBL; AY078808; AAL89528.1; JOINED.
DR   EMBL; AY078809; AAL89528.1; JOINED.
DR   EMBL; AY078810; AAL89528.1; JOINED.
DR   EMBL; AY078811; AAL89528.1; JOINED.
DR   EMBL; AY078812; AAL89528.1; JOINED.
DR   GO; GO:0005529; F:sugar binding; IEA.
DR   InterPro; IPR002353; AntifreezeII.
DR   InterPro; IPR001304; Lectin_C.
DR   Pfam; PF00059; lectin_c; 1.
DR   PRINTS; PR00356; ANTIFREEZEII.
DR   SMART; SM00034; CLECT; 1.
DR   PROSITE; PS00615; C_TYPE_LECTIN_1; 1.
DR   PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
SQ   SEQUENCE   399 AA;  45404 MW;  DEC669244205B27A CRC64;

  Query Match             13.5%;  Score 150.5;  DB 6;  Length 399;
  Best Local Similarity   32.4%;  Pred. No. 4.5e-06;
  Matches   48;  Conservative   19;  Mismatches   48;  Indels   33;  Gaps    9;

Qy         25 YFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFWIGLRRREE 84
              ||   |:| |: ::  |||  | ||| |:| :||  ::         |:   |:||   :
Db        280 YFISNSQR-NWHDSVTACREVGAQLVVIKSAEEQNFLQLQSSR---SNRFAWMGLSDLNQ 335

Qy         85 KQSNSTACQDLYAWTDGS--ISQF-RNWYVDEP-SCGSEVCVVMYHQPSAPAGIGGPYMF 140
                     :  : | |||   | | | |   || : | | |          :| |
Db        336 --------EGTWQWVDGSPLSSSFQRYWNSGEPNNSGDEDCAEF-------SGSG----- 375

Qy        141 QWNDDRCNMKNNFICKYSDEKPAVPSRE 168
              |||:|||: | :|||    ||    |:
Db        376 -WNDNRCNVDNYWICK----KPTACFRD 398


RESULT 47
Q28008
ID   Q28008      PRELIMINARY;      PRT;   197 AA.
AC   Q28008;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   C-type lectin homolog precursor.
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Cartilage;
RA   Neame P.J., Boynton R.E.;
RT   "C-type lectin homolog from bovine cartilage.";
RL   Submitted (APR-1998) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U22298; AAC18614.1; -.
DR   HSSP; P05452; 1TN3.
DR   GO; GO:0005529; F:sugar binding; IEA.
DR   GO; GO:0007157; P:heterophilic cell adhesion; IEA.
DR   InterPro; IPR001304; Lectin_C.
DR   Pfam; PF00059; lectin_c; 1.
DR   SMART; SM00034; CLECT; 1.
DR   PROSITE; PS00615; C_TYPE_LECTIN_1; 1.
DR   PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
KW   Lectin; Signal.
FT   SIGNAL        1     24       POTENTIAL.
FT   CHAIN        25    197       C-TYPE LECTIN HOMOLOG.
SQ   SEQUENCE   197 AA;  22215 MW;  AAAC4280F41AC0F4 CRC64;

  Query Match             13.5%;  Score 150;  DB 6;  Length 197;
  Best Local Similarity   25.0%;  Pred. No. 2.1e-06;
  Matches   38;  Conservative   24;  Mismatches   66;  Indels   24;  Gaps    4;

Qy          9 QPVCRGGTQ--RPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIE 66
              | ||  ||:  : ||             :| || | |   || ||   | ||   :  : :
Db         65 QTVCLRGTKFHKKCYLA-----AEGLKHFHEANEDCISKGGTLVVPRSADEINALRDYGK 119

Qy         67 NLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEPSCGSEVCVVMYH 126
               ||   |||:|:         |    :  :   :|    | ||   :|: |       ::
Db        120 RSLPGVNDFWLGI--------NDMVAEGKFVDINGLAISFLNWDQAQPNGGKRENCALFS 171

Qy        127 QPSAPAGIGGPYMFQWNDDRCNMKNNFICKYS 158
              | :           :|:|: |:    :||:::
Db        172 QSA---------QGKWSDEACHSSKRYICEFT 194


RESULT 48
Q8JZX6
ID   Q8JZX6      PRELIMINARY;      PRT;   158 AA.
AC   Q8JZX6;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   RIKEN cDNA 3110037K17 gene.
GN   3110037K17RIK.
OS   Mus musculus (Mouse).
```

```
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Mammary gland;
RA   Strausberg R.;
RL   Submitted (JUL-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC034893; AAH34893.1; -.
DR   MGD; MGI:1920399; 3110037K17Rik.
DR   GO; GO:0004190; F:aspartic-type endopeptidase activity; IEA.
DR   GO; GO:0005529; F:sugar binding; IEA.
DR   GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR   InterPro; IPR002353; AntifreezeII.
DR   InterPro; IPR001969; Aspprotease_AS.
DR   InterPro; IPR001304; Lectin_C.
DR   Pfam; PF00059; lectin_c; 1.
DR   PRINTS; PR00356; ANTIFREEZEII.
DR   SMART; SM00034; CLECT; 1.
DR   PROSITE; PS00141; ASP_PROTEASE; 1.
DR   PROSITE; PS00615; C_TYPE_LECTIN_1; 1.
DR   PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
SQ   SEQUENCE   158 AA;  18325 MW;  B34D8A47FDA4F76B CRC64;

  Query Match             13.4%;  Score 149.5;  DB 11;  Length 158;
  Best Local Similarity   29.3%;  Pred. No. 1.8e-06;
  Matches   41;  Conservative   21;  Mismatches   55;  Indels   23;  Gaps    4;

Qy          18 RPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFWI 77
               :|     ||  |    :: |:|| |   |  || | |::||     ||  :| :   ::|
Db          33 KPFGSYCYFTSTDLVASWNESKENCFHMGAHLVVIHSQEEQ----DFITGILDTGTAYFI 88

Qy          78 GLRRREEKQSNSTACQDLYAWTDGSISQFRN--WYVDEPSCGSEVCVVMYHQPSAPAGIG 135
               ||    ::|         : | | :        |:  |||  :| ||:: |: |    |
Db          89 GLSNPGDQQ---------WQWIDQTPYDDNTTFWHKGEPSSDNEQCVIINHRQSTGWG-- 137

Qy         136 GPYMFQWNDDRCNMKNNFIC 155
                 |:|  |: | | ||
Db         138 ------WSDIPCSDKQNSIC 151


RESULT 49
Q22136
ID   Q22136      PRELIMINARY;      PRT;   459 AA.
AC   Q22136;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   T04A8.3 protein.
GN   T04A8.3.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Palmer S.;
RL   Submitted (AUG-1994) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=99069613; PubMed=9851916;
RA   none;
RT   "Genome sequence of the nematode C.elegans: A platform for
RT   investigating biology.";
RL   Science 282:2012-2018(1998).
DR   EMBL; Z35663; CAA84726.1; -.
DR   PIR; T24425; T24425.
DR   WormPep; T04A8.3; CE01072.
DR   GO; GO:0005529; F:sugar binding; IEA.
DR   InterPro; IPR001304; Lectin_C.
DR   Pfam; PF00059; lectin_c; 1.
DR   SMART; SM00034; CLECT; 1.
DR   PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
SQ   SEQUENCE   459 AA;  53253 MW;  BD173368D31732A0 CRC64;

  Query Match             13.4%;  Score 149.5;  DB 5;  Length 459;
  Best Local Similarity   28.5%;  Pred. No. 6.6e-06;
  Matches   49;  Conservative   23;  Mismatches   52;  Indels   48;  Gaps    9;

Qy          20 CYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDF--WI 77
               |:|  | || |  :| ::|:: |   |  | | |  : :| ||::: |  : |:| :   |:
Db         287 CHK--YVHDMS--MNNDDAEKKCNDMGAHLSSFTTYEELKLLDEMILEVYPNDNNIAVWL 342

Qy          78 GLRRREEKQSNSTACQDL--------------------YAWTDGSISQ---------FR 107
               | :|||       | ||                    : | :| ::|         |
Db         343 GAKRREE-------CGDLSKNFTGGYSKDIHDPCARSRVFEWQNG-VAQNPPIFVGDGFD 394

Qy         108 NW---YVDEPSCGSEVCVV-MYHQPSAPAGIGGPYMFQWNDDRCNMKNNFIC 155
                 |   |     |  || |:| |    |   |    |   | ||  || :  |:|
Db         395 YWAEKYEPNHSTDSERCLVQMSGSMSVWYGDNKPRNMQINDIYCNYEFKFLC 446


RESULT 50
Q90WI8
ID   Q90WI8      PRELIMINARY;      PRT;   158 AA.
AC   Q90WI8;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   C-type lectin-like protein 1.
OS   Bungarus fasciatus (Banded krait).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
OC   Elapidae; Bungarinae; Bungarus.
OX   NCBI_TaxID=8613;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Venom gland;
RA   Zha H.-G., Zhang Y.;
RT   "cDNA cloning and characterization of C-type lectin-like proteins from
RT   Elapidae snakes.";
RL   Submitted (FEB-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF354270; AAK43584.1; -.
DR   GO; GO:0005529; F:sugar binding; IEA.
DR   GO; GO:0007157; P:heterophilic cell adhesion; IEA.
DR   InterPro; IPR001304; Lectin_C.
DR   InterPro; IPR003990; Pancreatis_ac.
DR   Pfam; PF00059; lectin_c; 1.
DR   PRINTS; PR01504; PNCREATITSAP.
DR   SMART; SM00034; CLECT; 1.
DR   PROSITE; PS00615; C_TYPE_LECTIN_1; 1.
DR   PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
KW   Lectin.
SQ   SEQUENCE   158 AA;  18638 MW;  ECF85936FA4182B8 CRC64;

  Query Match             13.4%;  Score 149;  DB 13;  Length 158;
  Best Local Similarity   28.4%;  Pred. No. 2e-06;
  Matches   40;  Conservative   20;  Mismatches   55;  Indels   26;  Gaps    5;

Qy          20 CYKVIYFHDTSRRLNFEEAKEACR--RDGGQLVSIESEDEQKLIEKFIENLLPSDGDFWI 77
               ||||   | |     : :|:  ||  : |  | |: |  :     ::| : |   |  ||
Db          37 CYKVFSKHKT-----WFDAEMYCRKFKPGCHLASLHSNADAVEFSEYISDYLTGQGHVWI 91

Qy          78 GLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEPS--CGSEVCVVMYHQPSAPAGIG 135
               |||  ::|         :: ||| | : |  |   :|     :| || :         :
Db          92 GLRDTKKKY--------IWEWTDRSRTDFLPWRKKQPDHFNNNEFCVEI---------VN 134

Qy         136 GPYMFQWNDDRCNMKNNFICK 156
                  ||||| |     |:|:
Db         135 FTGYLQWNDDNCAALRPFLCQ 155


Search completed: September  9, 2004, 22:38:52
Job time : 119 secs
```

OM protein - protein search, using sw model

Run on: September 9, 2004, 22:31:58 ; Search time 40 Seconds
(without alignments)
495.386 Million cell updates/sec

Title: US-09-887-855-2_COPY_22_227
Perfect score: 1115
Sequence: 1 ATGRLLSGQPVCRGGTQRPC.....EEDAKKTFKESREAALNLAY 206

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 50 summaries

Database : PIR_78:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	182	16.3	1456	1	A36563	mannose receptor p
2	178.5	16.0	1455	1	A48925	mannose receptor p
3	177	15.9	1268	2	S52781	neurocan - mouse
4	174.5	15.7	1643	2	T14274	versican precursor
5	174.5	15.7	3381	2	T42389	versican precursor
6	174	15.6	1257	2	S28764	neurocan precursor
7	174	15.6	2397	1	A55535	versican precursor
8	174	15.6	2409	1	A60979	versican precursor
9	171	15.3	3562	2	A47171	chondroitin sulfat
10	170.5	15.3	1479	2	T42710	mannose receptor,
11	158.5	14.2	1340	2	A39808	proteoglycan core
12	158.5	14.2	2327	2	T42630	aggrecan - bovine
13	158.5	14.2	2415	1	A39086	aggrecan precursor
14	154.5	13.9	612	2	B42755	E-selectin precurs
15	153.5	13.8	2124	2	A28452	proteoglycan core
16	152	13.6	912	2	A54423	brevican precursor
17	149.5	13.4	459	2	T24425	hypothetical prote
18	149	13.4	321	1	LNHUER	IgE Fc receptor II
19	148.5	13.3	330	2	T46256	brevican - human (
20	148.5	13.3	2132	1	A55182	aggrecan precursor
21	148	13.3	253	2	E89130	protein F52E1.2 [i
22	147	13.2	883	2	S57653	brevican precursor
23	146.5	13.1	162	1	LNRC1	lectin BRA3-1 prec
24	146	13.1	2109	1	I50421	aggrecan precursor
25	145.5	13.0	742	2	JC7595	scavenger receptor
26	145	13.0	883	2	S49126	brevican precursor
27	144.5	13.0	173	2	S10548	lectin - barnacle
28	144.5	13.0	372	2	S23936	L-selectin precurs
29	144.5	13.0	404	2	A46274	HIV gp120-binding
30	143.5	12.9	129	2	JC4329	coagulation factor
31	143.5	12.9	372	1	A32375	L-selectin precurs
32	143.5	12.9	463	2	T26655	hypothetical prote
33	142.5	12.8	131	2	JC5058	bitiscetin alpha c
34	142	12.7	1487	2	S48719	phospholipase-A(2)
35	141.5	12.7	331	1	LNMSER	IgE Fc receptor, l
36	140.5	12.6	162	1	LNRC3	lectin BRA3-2 prec
37	140	12.6	370	2	S22124	L-selectin precurs
38	139	12.5	248	1	LNHUPS	pulmonary surfacta
39	139	12.5	248	1	LNHUP6	pulmonary surfacta
40	139	12.5	248	1	LNHUP1	pulmonary surfacta
41	139	12.5	283	1	LNFHLS	lectin precursor -
42	138.5	12.4	152	2	JC4690	coagulation factor
43	138.5	12.4	202	2	JC4031	tetranectin precur
44	138	12.4	280	2	T29200	hypothetical prote
45	137.5	12.3	309	1	S34198	IgE Fc receptor II
46	137.5	12.3	372	2	JC5377	L-selectin precurs
47	137.5	12.3	376	2	JC4892	L-selectin precurs
48	137	12.3	248	2	I51921	pulmonary surfacta
49	136.5	12.2	308	2	T29754	hypothetical prote
50	136	12.2	175	2	A37194	pancreatic thread

ALIGNMENTS

RESULT 1
A36563
mannose receptor precursor - human
C;Species: Homo sapiens (man)
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C;Accession: A36563; A60926; A44255; B44255; C44255; D44255; E44255; F44255; G44255; H44
R;Taylor, M.E.; Conary, J.T.; Lennartz, M.R.; Stahl, P.D.; Drickamer, K.
J. Biol. Chem. 265, 12156-12162, 1990
A;Title: Primary structure of the mannose receptor contains multiple motifs resembling c
A;Reference number: A36563; MUID:90324192; PMID:2373685
A;Accession: A36563
A;Molecule type: mRNA
A;Residues: 1-1456 <TAY>
A;Cross-references: GB:J05550; NID:g188675; PIDN:AAA59868.1; PID:g188676
A;Note: parts of this sequence, including the amino end of the mature protein, were conf
R;Ezekowitz, R.A.B.; Sastry, K.; Bailly, P.; Warner, A.
J. Exp. Med. 172, 1785-1794, 1990
A;Title: Molecular characterization of the human macrophage mannose receptor: demonstrat
A;Reference number: A60926; MUID:91079783; PMID:2258707
A;Accession: A60926
A;Status: nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1-1333,'T',1335-1456 <EZE>
A;Cross-references: GB:X55635
A;Note: translation of the nucleotide sequence is incomplete
A;Note: in the authors' translation additional residues Pro-Glu-Ile are shown after 497-
R;Kim, S.J.; Ruiz, N.; Bezouska, K.; Drickamer, K.
Genomics 14, 721-727, 1992
A;Title: Organization of the gene encoding the human macrophage mannose receptor (MRC1).
A;Reference number: A44255; MUID:93052405; PMID:1294118
A;Accession: A44255
A;Status: preliminary; nucleic acid sequence not shown; not compared with conceptual tra
A;Molecule type: DNA
A;Residues: 155-233,'KSAL',238-283;346-428;492-569;631-714,716-719;783-820,'N',822-865;'
A;Note: sequence extracted from NCBI backbone (NCBIP:118415, NCBIP:118421, NCBIP:118428,
C;Genetics:
A;Gene: GDB:MRC1
A;Cross-references: GDB:133759; OMIM:153618
A;Map position: 10p13-10p13
C;Superfamily: phospholipase A2 receptor; C-type lectin homology; fibronectin type II re
C;Keywords: duplication; lectin; tandem repeat; transmembrane protein
F;1-18/Domain: signal sequence #status predicted <SIG>
F;168-209/Domain: fibronectin type II repeat homology <2F1>
F;223-340/Domain: C-type lectin homology <LCH1>
F;362-486/Domain: C-type lectin homology <LCH2>
F;945-1079/Domain: C-type lectin homology <LCH3>

```
  Query Match             16.3%;  Score 182;  DB 1;  Length 1456;
  Best Local Similarity   25.5%;  Pred. No. 9.1e-08;
  Matches   52;  Conservative   37;  Mismatches   61;  Indels   54;  Gaps    8;

Qy         21 YKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFWIGLR 80
              ||   :: :  :   : |:  |:|: | ||||:|| |:| : |:: |   :   ::|||
Db        807 YKDYQYYFSKEKETMDNARAFCKRNFGDLVSIQSESEKKFLWKYV-NRNDAQSAYFIGLL 865

Qy         81 RREEKQSNSTACQDLYAWTDGSISQFRNWYVDEPSCGS--EVCVVMYHQPSAPAGIGGPY 138
                :|:         :|| |||   : :|   ||:  :  | || ||    : :|
Db        866 ISLDKK---------FAWMDGSKVDYVSWATGEPNFANEDENCVTMY----SNSGF---- 908

Qy        139 MFQWNDDRCNMKNNFICKYSDEK----PAVPSREAEGEETELTTPVLPEETQE------- 187
               |||  |   | |||:  :         :|           | | :|   :|
Db        909 ---WNDINCGYPNAFICQRHNSSINATTVMP----------TMPSVPSGCKEGWNFYSN 954

Qy        188 ---------EDAKKTFKESREAAL 202
                      |: :| ::|:|:| :
Db        955 KCFKIFGFMEEERKNWQEARKACI 978
```

RESULT 2
A48925
mannose receptor precursor, macrophage - mouse
C;Species: Mus musculus (house mouse)
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C;Accession: A48925; S21320; PC2245
R;Harris, N.; Super, M.; Rits, M.; Chang, G.; Ezekowitz, R.A.
Blood 80, 2363-2373, 1992
A;Title: Characterization of the murine macrophage mannose receptor: demonstration that
A;Reference number: A48925; MUID:93043353; PMID:1421407
A;Accession: A48925
A;Status: not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-1455 <HAR>
A;Experimental source: peritoneal macrophage
A;Note: sequence extracted from NCBI backbone (NCBIP:118733)
R;Harris, N.; Super, M.; Rits, M.; Chang, G.; Ezekowitz, R.B.
submitted to the EMBL Data Library, April 1992
A;Description: Characterization of the murine macrophage mannose receptor: Demonstration
on.
A;Reference number: S21320
A;Accession: S21320
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-302,'W',303-1117,'E',1119-1455 <HA2>
A;Cross-references: EMBL:Z11974; NID:g52997; PIDN:CAA78028.1; PID:g52998
R;Harris, N.; Peters, L.L.; Eicher, E.M.; Rits, M.; Raspberry, D.; Eichbaum, Q.G.; Super
Biochem. Biophys. Res. Commun. 198, 682-692, 1994
A;Title: The exon-intron structure and chromosomal localization of the mouse macrophage
A;Reference number: PC2245; MUID:94128116; PMID:8297379
A;Accession: PC2245
A;Molecule type: mRNA
A;Residues: 35-105 <HA3>
C;Genetics:
A;Gene: Mrc1
A;Map position: 2
C;Superfamily: phospholipase A2 receptor; C-type lectin homology; fibronectin type II re
C;Keywords: membrane protein; receptor
F;168-209/Domain: fibronectin type II repeat homology <2F9>
F;361-485/Domain: C-type lectin homology <LCH1>
F;943-1077/Domain: C-type lectin homology <LCH2>

```
  Query Match             16.0%;  Score 178.5;  DB 1;  Length 1455;
  Best Local Similarity   25.9%;  Pred. No. 1.9e-07;
  Matches   53;  Conservative   33;  Mismatches   66;  Indels   53;  Gaps    8;

Qy         21 YKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFWIGLR 80
              ||   :: :  :   : |:  |::: | | :|:|| |:| : |:| |       ::||:
Db        806 YKDYQYYFSKEKETMDNARRFCKKNFGDLATIKSESEKKFLWKYI-NKNGGQSPYFIGML 864

Qy         81 RREEKQSNSTACQDLYAWTDGSISQFRNWYVDEPSCGS--EVCVVMYHQPSAPAGIGGPY 138
                :|:         : | |||   |  |   ||:  :  | || ||      :|
```

```
Db        865 ISMDKK---------FIWMDGSKVDFVAWATGEPNFANDDENCVTMY----TNSGF---- 907

Qy        139 MFQWNDDRCNMKNNFICK---YSDEKPAVPSREAEGEETELTTPVLPEETQE-------- 187
                 |||  |   |||||:    |    |:|           |||  |   :|
Db        908 ---WNDINCGYPNNFICQRHNSSINATAMP----------TTPTTPGGCKEGWHLYKNK 953

Qy        188 --------EDAKKTFKESREAALNL 204
                      : ||:::::|:|   |
Db        954 CFKIFGFANEEKKSWQDARQACKGL 978
```

RESULT 3
S52781
neurocan - mouse
C;Species: Mus musculus (house mouse)
C;Date: 19-May-1995 #sequence_revision 21-Jul-1995 #text_change 04-Feb-2000
C;Accession: S52781
R;Rauch, U.; Forsberg, N.; Kulbe, G.; Arnold-Ammer, I.; Faessler, R.
submitted to the EMBL Data Library, February 1995
A;Description: Amino acid sequence of mouse neurocan and brevican and their different exp
A;Reference number: S52781
A;Accession: S52781
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-1268 <RAU>
A;Cross-references: EMBL:X84727; NID:g758629; PIDN:CAA59216.1; PID:g758630
C;Superfamily: aggrecan; C-type lectin homology; complement factor H repeat homology; EGF
F;176-253/Domain: link protein repeat homology <LNK1>
F;274-355/Domain: link protein repeat homology <LNK2>
F;964-995/Domain: EGF homology <EGF>
F;1040-1160/Domain: C-type lectin homology <LCH>
F;1167-1223/Domain: complement factor H repeat homology <FHD>

```
  Query Match             15.9%;  Score 177;  DB 2;  Length 1268;
  Best Local Similarity   31.5%;  Pred. No. 2.2e-07;
  Matches   45;  Conservative   17;  Mismatches   49;  Indels   32;  Gaps    6;

Qy         17 QRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFW 76
              |  ||:  ||   : |  :|:|:  |||  | | |: | :| | |  |         : |
Db       1048 QGHCYR--YF---AHRRAWEDAERDCRRRAGHLTSVHSPEEHKFINSF------GHENSW 1096

Qy         77 IGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEPS---CGSEVCVVMYHQPSAPAG 133
              |||  |  ::         : ||| :  |: ||   :|     | | ||||    |
Db       1097 IGLNDRTVERD--------FQWTDNTGLQYENWREKQPDNFFAGGEDCVVMVAHESG--- 1145

Qy        134 IGGPYMFQWNDDRCNMKNNFICK 156
                    :|||  ||    ::||
Db       1146 -------RWNDVPCNYNLPYVCK 1161
```

RESULT 4
T14274
versican precursor, splice form V2 - bovine
C;Species: Bos primigenius taurus (cattle)
C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 05-May-2000
C;Accession: T14274
R;Schmalfeldt, M.; Dours-Zimmermann, M.T.; Winterhalter, K.H.; Zimmermann, D.R.
J. Biol. Chem. 273, 15758-15764, 1998
A;Title: Versican V2 is a major extracellular matrix component of the mature bovine brain
A;Reference number: Z17954; MUID:98288320; PMID:9624174
A;Accession: T14274
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-1643 <SCH>
A;Cross-references: EMBL:AF060458; NID:g3253303; PID:g3253304; PIDN:AAC24360.1
A;Experimental source: brain
C;Keywords: glycoprotein
F;1-20/Domain: signal sequence #status predicted <SIG>
F;21-1643/Product: versican, splice form V2 #status predicted <MAT>
F;57,331,352,817,965,1017,1333,1616,1626/Binding site: carbohydrate (Asn) (covalent) #sta

```
  Query Match             15.7%;  Score 174.5;  DB 2;  Length 1643;
  Best Local Similarity   25.5%;  Pred. No. 4.8e-07;
```

```
  Matches   59;  Conservative   26;  Mismatches   71;  Indels   75;  Gaps   11;

Qy         17 QRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDF- 75
              |  |||  ||   : |  :: |:  ||  |  | || | :||  : :          |:
Db       1424 QGQCYK--YF---AHRRTWDAAERECRLQGAHLTSILSHEEQMFVNRV-------GHDYQ 1471

Qy         76 WIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEP----SCGSEVCVVMYHQPSAP 131
              ||||        |    :  : |||||  |: ||  ::|    | | :  |:::|:
Db       1472 WIGL--------NDKMFEHDFRWTDGSTLQYENWRPNQPDSFFSTGEDCVVIIWHENG-- 1521

Qy        132 AGIGGPYMFQWNDDRCNMKNNFICKYS----DEKPAVPSREAEGE-------------- 172
                  ||||  ||    : ||       : | | : :  |:
Db       1522 ---------QWNDVPCNYHLTYTCKKGTVACGQPPVVENAKTFGKMKPRYEINSLIRYHC 1572

Qy        173 -----ETELTT----------PVL----PEETQEEDAKKTFKESREAALN 203
                :  | |            | :    |   |   :|| || | |  |
Db       1573 KDGFIQRHLPTIRCLGNGRWAMPKITCLNPSAYQRTYSKKYFKNSSSAKDN 1623


RESULT 5
T42389
versican precursor, splice form V0 - bovine
N;Alternate names: chondroitin sulfate proteoglycan
C;Species: Bos primigenius taurus (cattle)
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 05-May-2000
C;Accession: T42389
R;Schmalfeldt, M.; Dours-Zimmermann, M.T.; Winterhalter, K.H.; Zimmermann, D.R.
J. Biol. Chem. 273, 15758-15764, 1998
A;Title: Versican V2 is a major extracellular matrix component of the mature bovine brai
A;Reference number: Z17954; MUID:98288320; PMID:9624174
A;Accession: T42389
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-3381 <SCH>
A;Cross-references: EMBL:AF060456; NID:g3253299; PID:g3253300; PIDN:AAC24358.1
C;Superfamily: chicken chondroitin sulfate proteoglycan PG-M core protein; C-type lectin
C;Keywords: chondroitin sulfate proteoglycan; extracellular matrix; glycoprotein
F;1-20/Domain: signal sequence #status predicted <SIG>
F;21-3381/Product: versican, splice form V0 #status predicted <MAT>
F;57,331,352,817,965,1017,1333,1393,1437,1463,1653,1974,2045,2074,2103,2263,2290,2356,26

  Query Match             15.7%;  Score 174.5;  DB 2;  Length 3381;
  Best Local Similarity   25.5%;  Pred. No. 1.1e-06;
  Matches   59;  Conservative   26;  Mismatches   71;  Indels   75;  Gaps   11;

Qy         17 QRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDF- 75
              |  |||  ||   : |  :: |:  ||  |  | || | :||  : :          |:
Db       3162 QGQCYK--YF---AHRRTWDAAERECRLQGAHLTSILSHEEQMFVNRV-------GHDYQ 3209

Qy         76 WIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEP----SCGSEVCVVMYHQPSAP 131
              ||||        |    :  : |||||  |: ||  ::|    | | :  |:::|:
Db       3210 WIGL--------NDKMFEHDFRWTDGSTLQYENWRPNQPDSFFSTGEDCVVIIWHENG-- 3259

Qy        132 AGIGGPYMFQWNDDRCNMKNNFICKYS----DEKPAVPSREAEGE-------------- 172
                  ||||  ||    : ||       : | | : :  |:
Db       3260 ---------QWNDVPCNYHLTYTCKKGTVACGQPPVVENAKTFGKMKPRYEINSLIRYHC 3310

Qy        173 -----ETELTT----------PVL----PEETQEEDAKKTFKESREAALN 203
                :  | |            | :    |   |   :|| || | |  |
Db       3311 KDGFIQRHLPTIRCLGNGRWAMPKITCLNPSAYQRTYSKKYFKNSSSAKDN 3361


RESULT 6
S28764
neurocan precursor - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 22-Nov-1993 #sequence_revision 01-Sep-1995 #text_change 04-Feb-2000
C;Accession: S28764
R;Rauch, U.; Karthikeyan, L.; Maurel, P.; Margolis, R.U.; Margolis, R.K.
J. Biol. Chem. 267, 19536-19547, 1992
A;Title: Cloning and primary structure of neurocan, a developmentally regulated, aggrega
A;Reference number: S28764; MUID:92406907; PMID:1326557
A;Accession: S28764
A;Molecule type: mRNA
A;Residues: 1-1257 <RAU>
A;Cross-references: EMBL:M97161; NID:g205649; PIDN:AAC37679.1; PID:g205650
C;Superfamily: aggrecan; C-type lectin homology; complement factor H repeat homology; EG
C;Keywords: chondroitin sulfate proteoglycan; glycoprotein
F;1-22/Domain: signal sequence #status predicted <SIG>
F;23-1257/Product: neurocan #status predicted <MAT>
F;176-253/Domain: link protein repeat homology <LNK1>
F;274-355/Domain: link protein repeat homology <LNK2>
F;364-366/Region: cell attachment (R-G-D) motif
F;953-984/Domain: EGF homology <EGF>
F;1029-1149/Domain: C-type lectin homology <LCH>
F;1156-1212/Domain: complement factor H repeat homology <FHD>
F;121,339,737,967,1164/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;372,410/Binding site: chondroitin sulfate (Ser) (covalent) #status predicted
F;944/Binding site: chondroitin sulfate (Ser) (covalent) #status experimental

  Query Match             15.6%;  Score 174;  DB 2;  Length 1257;
  Best Local Similarity   30.8%;  Pred. No. 3.9e-07;
  Matches   44;  Conservative   18;  Mismatches   49;  Indels   32;  Gaps    6;

Qy         17 QRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFW 76
              |  ||:  ||   : |  :|:|:  |||  | | |: | :| | |  |        : |
Db       1037 QGHCYR--YF---AHRRAWEDAERDCRRRAGHLTSVHSPEEHKFINSF------GHENSW 1085

Qy         77 IGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEPS---CGSEVCVVMYHQPSAPAG 133
              |||  |  ::         : |||  :  |: ||   :|     | | ||||    :
Db       1086 IGLNDRTVERD--------FQWTDNTGLQYENWREKQPDNFFAGGEDCVVMVAHENG--- 1134

Qy        134 IGGPYMFQWNDDRCNMKNNFICK 156
                  :|||  ||    ::||
Db       1135 -------RWNDVPCNYNLPYVCK 1150


RESULT 7
A55535
versican precursor - mouse
N;Alternate names: chondroitin sulfate proteoglycan 2; chondroitin sulfate proteoglycan
versican
N;Contains: glial hyaluronate-binding protein
C;Species: Mus musculus (house mouse)
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C;Accession: A55535
R;Ito, K.; Shinomura, T.; Zako, M.; Ujita, M.; Kimata, K.
J. Biol. Chem. 270, 958-965, 1995
A;Title: Multiple forms of mouse PG-M, a large chondroitin sulfate proteoglycan generate
A;Reference number: A55535; MUID:95122551; PMID:7822336
A;Accession: A55535
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-2397 <RES>
A;Cross-references: GB:D16263; NID:g862460; PIDN:BAA03796.1; PID:g862461
C;Superfamily: versican; C-type lectin homology; complement factor H repeat homology; EG
F;1-20/Domain: signal sequence #status predicted <SIG>
F;21-1654/Domain: versican #status predicted <MAT>
F;167-244/Domain: link protein repeat homology <LNK1>
F;265-346/Domain: link protein repeat homology <LNK2>
F;2095-2126/Domain: EGF homology <EG1>
F;2133-2164/Domain: EGF homology <EG2>
F;2171-2291/Domain: C-type lectin homology <LCH>
F;2298-2354/Domain: complement factor H repeat homology <FHD>

  Query Match             15.6%;  Score 174;  DB 1;  Length 2397;
  Best Local Similarity   28.5%;  Pred. No. 8.3e-07;
  Matches   47;  Conservative   23;  Mismatches   55;  Indels   40;  Gaps    8;

Qy         17 QRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDF- 75
              |  |||  ||   : |  :: |:  ||  |  | || | :||  : :          |:
Db       2179 QGQCYK--YF---AHRRTWDAAERECRLQGAHLTSILSHEEQMFVNRV-------GHDYQ 2226

Qy         76 WIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEP----SCGSEVCVVMYHQPSAP 131
              ||||        |    :  : |||||  |: ||  ::|    | | :  |:::|:
Db       2227 WIGL--------NDKMFEHDFRWTDGSALQYENWRPNQPDSFFSAGEDCVVIIWHENG-- 2276
```

```
Qy        132 AGIGGPYMFQWNDDRCNMKNNFICKYS----DEKPAVPSREAEGE 172
                      ||||  ||    : ||       : | | : :  |:
Db       2277 ---------QWNDVPCNYHLTYTCKKGTVACGQPPVVENAKTFGK 2312
```

RESULT 8
A60979
versican precursor - human
N;Alternate names: chondroitin sulfate proteoglycan 2; chondroitin sulfate proteoglycan
N;Contains: glial hyaluronate-binding protein
C;Species: Homo sapiens (man)
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 19-Jan-2001
C;Accession: S06014; S43921; A60979; A30358; A29348; A45131; I54179
R;Zimmermann, D.R.; Ruoslahti, E.
EMBO J. 8, 2975-2981, 1989
A;Title: Multiple domains of the large fibroblast proteoglycan, versican.
A;Reference number: S06014; MUID:90059882; PMID:2583089
A;Accession: S06014
A;Molecule type: mRNA
A;Residues: 1-2409 <ZIM>
A;Cross-references: GB:X15998; NID:g37662; PIDN:CAA34128.1; PID:g37663
R;Yao, L.Y.; Moody, C.; Schoenherr, E.; Wight, T.N.; Sandell, L.J.
Matrix Biol. 14, 213-225, 1994
A;Title: Identification of the proteoglycan versican in aorta and smooth muscle cells by
A;Reference number: S43921; MUID:95005762; PMID:7921538
A;Accession: S43921
A;Molecule type: mRNA
A;Residues: 208-440;1094-1385;1910-2246 <YAO>
R;Bignami, A.; Lane, W.S.; Andrews, D.; Dahl, D.
Brain Res. Bull. 22, 67-70, 1989
A;Title: Structural similarity of hyaluronate binding proteins in brain and cartilage.
A;Reference number: A60979; MUID:89229983; PMID:2469524
A;Accession: A60979
A;Molecule type: protein
A;Residues: 171-210;289-303 <BIG>
R;Perides, G.; Lane, W.S.; Andrews, D.; Dahl, D.; Bignami, A.
J. Biol. Chem. 264, 5981-5987, 1989
A;Title: Isolation and partial characterization of a glial hyaluronate-binding protein.
A;Reference number: A30358; MUID:89174663; PMID:2466833
A;Accession: A30358
A;Molecule type: protein
A;Residues: 24-50;80-87,'D',89-119;128-155;167-218;229-259,'IR';261-268;277-283,'G',285-
R;Krusius, T.; Gehlsen, K.R.; Ruoslahti, E.
J. Biol. Chem. 262, 13120-13125, 1987
A;Title: A fibroblast chondroitin sulfate proteoglycan core protein contains lectin-like
A;Reference number: A29348; MUID:88007514; PMID:2820964
A;Accession: A29348
A;Molecule type: mRNA
A;Residues: 1725,'V',1727-2409 <KRU>
A;Cross-references: GB:J02814
R;Perides, G.; Rahemtulla, F.; Lane, W.S.; Asher, R.A.; Bignami, A.
J. Biol. Chem. 267, 23883-23887, 1992
A;Title: Isolation of a large aggregating proteoglycan from human brain.
A;Reference number: A45131; MUID:93054750; PMID:1429726
A;Contents: brain
A;Accession: A45131
A;Molecule type: protein
A;Residues: 21-22,'X',24-37 <PE2>
A;Experimental source: brain
A;Note: sequence extracted from NCBI backbone (NCBIP:118884)
R;Iozzo, R.V.; Naso, M.F.; Cannizzaro, L.A.; Wasmuth, J.J.; McPherson, J.D.
Genomics 14, 845-851, 1992
A;Title: Mapping of the versican proteoglycan gene (CSPG2) to the long arm of human chrc
A;Reference number: I54179; MUID:93122792; PMID:1478664
A;Accession: I54179
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 251-347 <RES>
A;Cross-references: GB:S52488; NID:g263313; PIDN:AAB24878.1; PID:g263314
C;Genetics:
A;Gene: GDB:CSPG2
A;Cross-references: GDB:127873; OMIM:118661
A;Map position: 5q12-5q14
C;Superfamily: versican; C-type lectin homology; complement factor H repeat homology; EGI
F;1-20/Domain: signal sequence #status predicted <SIG>
F;21-2409/Product: proteoglycan 24K core protein #status predicted <MAT>
F;167-244/Domain: link protein repeat homology <LNK1>
F;265-346/Domain: link protein repeat homology <LNK2>
F;559-1654/Domain: chondroitin sulfate attachment #status predicted <GAG>
F;2106-2137/Domain: EGF homology <EG1>
F;2144-2175/Domain: EGF homology <EG2>
F;2182-2302/Domain: C-type lectin homology <LCH>
F;2309-2365/Domain: complement factor H repeat homology <FHD>

Query Match 15.6%; Score 174; DB 1; Length 2409;
Best Local Similarity 28.5%; Pred. No. 8.4e-07;
Matches 47; Conservative 23; Mismatches 55; Indels 40; Gaps 8;

```
Qy         17 QRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDF- 75
              |  |||  ||   : |  :: |:  ||  |  | || | :||  : :          |:
Db       2190 QGQCYK--YF---AHRRTWDAAERECRLQGAHLTSILSHEEQMFVNRV-------GHDYQ 2237

Qy         76 WIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEP----SCGSEVCVVMYHQPSAP 131
              ||||        |    :  : |||||  |: ||  ::|    | | :  |:::|:
Db       2238 WIGL--------NDKMFEHDFRWTDGSTLQYENWRPNQPDSFFSAGEDCVVIIWHENG-- 2287

Qy        132 AGIGGPYMFQWNDDRCNMKNNFICKYS----DEKPAVPSREAEGE 172
                     ||||  ||    : ||       : | | : :  |:
Db       2288 ---------QWNDVPCNYHLTYTCKKGTVACGQPPVVENAKTFGK 2323
```

RESULT 9
A47171
chondroitin sulfate proteoglycan PG-M core protein - chicken
C;Species: Gallus gallus (chicken)
C;Date: 21-Sep-1993 #sequence_revision 18-Nov-1994 #text_change 21-Jul-2000
C;Accession: A47171
R;Shinomura, T.; Nishida, Y.; Ito, K.; Kimata, K.
J. Biol. Chem. 268, 14461-14469, 1993
A;Title: cDNA cloning of PG-M, a large chondroitin sulfate proteoglycan expressed during

A;Reference number: A47171; MUID:93300846; PMID:8314802
A;Accession: A47171
A;Status: preliminary
A;Molecule type: nucleic acid
A;Residues: 1-3562 <SHI>
A;Cross-references: GB:D13542; NID:g391643; PIDN:BAA02742.1; PID:g391644
A;Experimental source: stage 22-23 developing limb buds
A;Note: sequence extracted from NCBI backbone (NCBIN:134456, NCBIP:134457)
C;Superfamily: chicken chondroitin sulfate proteoglycan PG-M core protein; C-type lectin
F;166-243/Domain: link protein repeat homology <LNK1>
F;264-345/Domain: link protein repeat homology <LNK2>
F;3258-3289/Domain: EGF homology <EGF1>
F;3296-3327/Domain: EGF homology <EGF>
F;3334-3454/Domain: C-type lectin homology <LCH>
F;3461-3517/Domain: complement factor H repeat homology <FHD>

Query Match 15.3%; Score 171; DB 2; Length 3562;
Best Local Similarity 28.5%; Pred. No. 2.4e-06;
Matches 47; Conservative 23; Mismatches 55; Indels 40; Gaps 8;

```
Qy         17 QRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDF- 75
              |  |||  ||   : |  :: |:  ||  |  | || | :||  : :          |:
Db       3342 QGQCYK--YF---AHRRTWDTAERECRLQGAHLTSILSHEEQVFVNRI-------GHDYQ 3389

Qy         76 WIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEP----SCGSEVCVVMYHQPSAP 131
              ||||        |    :  : |||||  |: ||  ::|    | | :  |:::|:
Db       3390 WIGL--------NDKMFERDFRWTDGSPLQYENWRPNQPDSFFSAGEDCVVIIWHENG-- 3439

Qy        132 AGIGGPYMFQWNDDRCNMKNNFICKYS----DEKPAVPSREAEGE 172
                     ||||  ||    : ||       : | | : :  |:
Db       3440 ---------QWNDVPCNYHLTYTCKKGTVACGQPPVVENAKTFGK 3475
```

RESULT 10

```
T42710
mannose receptor, macrophage - mouse
N;Alternate names: lambda lectin; phospholipase A2 receptor
C;Species: Mus musculus (house mouse)
C;Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 09-Jun-2000
C;Accession: T42710
R;Wu, K.; Yuan, J.; Lasky, L.A.
J. Biol. Chem. 271, 21323-21330, 1996
A;Title: Characterization of a novel member of the macrophage mannose receptor type C le
A;Reference number: Z22235; MUID:96355501; PMID:8702911
A;Accession: T42710
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-1479 <WUK>
A;Cross-references: EMBL:U56734; NID:g1336073; PID:g1336074; PIDN:AAC52729.1
C;Superfamily: phospholipase A2 receptor; C-type lectin homology; fibronectin type II re
C;Keywords: membrane protein; receptor
F;186-227/Domain: fibronectin type II repeat homology <2FR>

  Query Match             15.3%;  Score 170.5;  DB 2;  Length 1479;
  Best Local Similarity   31.4%;  Pred. No. 9.7e-07;
  Matches   48;  Conservative   22;  Mismatches   48;  Indels   35;  Gaps    7;

Qy          7 SGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIE 66
              | ||     |  ||::      : : :::|:| || | || |:|| |  | : | | |:
Db        384 SWQPF-----QGHCYRL-----QAEKRSWQESKRACLRGGGDLLSIHSMAELEFITKQIK 433

Qy         67 NLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEPS---CGSEVCVV 123
                :    : ||||        |    |  : |:|||:  | :|:  ||:      | ||
Db        434 QEVE---ELWIGL--------NDLKLQMNFEWSDGSLVSFTHWHPFEPNNFRDSLEDCVT 482

Qy        124 MYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICK 156
              ::     | |       :|||  ||     |||
Db        483 IW----GPEG-------RWNDSPCNQSLPSICK 504


RESULT 11
A39808
proteoglycan core protein, cartilage - bovine (fragments)
N;Alternate names: aggrecan; aggregating cartilage proteoglycan
C;Species: Bos primigenius taurus (cattle)
C;Date: 20-Mar-1992 #sequence_revision 23-Mar-1995 #text_change 13-Aug-1999
C;Accession: A34234; A27752; A39808; A27751; E29164; B27751; C27751; D27751; E27751; F27
R;Antonsson, P.; Heinegard, D.; Oldberg, A.
J. Biol. Chem. 264, 16170-16173, 1989
A;Title: The keratan sulfate-enriched region of bovine cartilage proteoglycan consists c
A;Reference number: A34234; MUID:89380219; PMID:2528543
A;Accession: A34234
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 128-621 <ANT>
A;Cross-references: GB:J05028
R;Oldberg, A.; Antonsson, P.; Heinegard, D.
Biochem. J. 243, 255-259, 1987
A;Title: The partial amino acid sequence of bovine cartilage proteoglycan,deduced from a
A;Reference number: A27752; MUID:87270630; PMID:3111460
A;Accession: A27752
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 622-1340 <OLD>
R;Sandy, J.D.; Boynton, R.E.; Flannery, C.R.
J. Biol. Chem. 266, 8198-8205, 1991
A;Title: Analysis of the catabolism of aggrecan in cartilage explants by quantitation of
A;Reference number: A39808; MUID:91217051; PMID:2022637
A;Accession: A39808
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-28;59-82;131-137,'QSET',142-149;196-207;226-249;1137-1143;1252-1267;1274-1
R;Perin, J.P.; Bonnet, F.; Jolles, P.
FEBS Lett. 206, 73-77, 1986
A;Title: Structural relationship between link proteins and proteoglycan monomers.
A;Reference number: A27751; MUID:87005253; PMID:3530809
A;Accession: A27751
A;Molecule type: protein
A;Residues: 29-58;74-130;174-175,'A',177-204;208-225 <PER>
R;Perin, J.P.; Bonnet, F.; Jolles, J.; Jolles, P.
FEBS Lett. 176, 37-42, 1984
A;Title: Sequence data concerning the protein core of the cartilage proteoglycan monomer
A;Reference number: A91327; MUID:85027710; PMID:6489519
A;Accession: E29164
A;Molecule type: protein
A;Residues: 1230-1249 <PE2>
C;Superfamily: aggrecan; C-type lectin homology; complement factor H repeat homology; EG
C;Keywords: glycoprotein
F;8-28/Domain: link protein repeat homology (fragment) <LNK1>
F;29-58/Domain: link protein repeat homology (fragment) <LNK2>
F;80-146/Domain: link protein repeat homology (fragments) <LNK3>
F;167-248/Domain: link protein repeat homology <LNK4>
F;1130-1250/Domain: C-type lectin homology <LCH>
F;1257-1313/Domain: complement factor H repeat homology <FHD>

  Query Match             14.2%;  Score 158.5;  DB 2;  Length 1340;
  Best Local Similarity   26.8%;  Pred. No. 1e-05;
  Matches   48;  Conservative   29;  Mismatches   59;  Indels   43;  Gaps   10;

Qy          9 QPVCRGG---TQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFI 65
              | :|  |     |  ||:  :| |   |  : :|:  ||:     | || : :||    :|:
Db       1127 QKLCEEGWTKFQGHCYR--HFPD---RATWVDAESQCRKQQSHLSSIVTPEEQ----EFV 1177

Qy         66 ENLLPSDGDF-WIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEP----SCGSEV 120
              |   :  |: ||||        |    :  : |:||   || ||  ::|    : | :
Db       1178 NN---NAQDYQWIGL--------NDKTIEGDFRWSDGHSLQFENWRPNQPDNFFATGEDC 1226

Qy        121 CVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYS----DEKPAVPSREAEGEETE 175
              |:::|:             :|||  || :  | ||       | | |      |:: :
Db       1227 VVMIWHEKG----------EWNDVPCNYQLPFTCKKGTVACGEPPVVEHARIFGQKKD 1274


RESULT 12
T42630
aggrecan - bovine
C;Species: Bos primigenius taurus (cattle)
C;Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 04-Mar-2000
C;Accession: T42630
R;Hering, T.M.; Kollar, J.; Huynh, T.D.
submitted to the EMBL Data Library, September 1996
A;Description: Complete coding sequence of bovine aggrecan: comparative structural analy
A;Reference number: Z22182
A;Accession: T42630
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-2327 <HER>
A;Cross-references: EMBL:U76615; NID:g1730259; PID:g1730260; PIDN:AAB38524.1
A;Experimental source: articular chondrocytes
C;Superfamily: aggrecan; C-type lectin homology; complement factor H repeat homology; EG
C;Keywords: cartilage; chondroitin sulfate proteoglycan; extracellular matrix; glycoprot

  Query Match             14.2%;  Score 158.5;  DB 2;  Length 2327;
  Best Local Similarity   26.8%;  Pred. No. 1.9e-05;
  Matches   48;  Conservative   29;  Mismatches   59;  Indels   43;  Gaps   10;

Qy          9 QPVCRGG---TQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFI 65
              | :|  |     |  ||:  :| |   |  : :|:  ||:     | || : :||    :|:
Db       2114 QKLCEEGWTKFQGHCYR--HFPD---RATWVDAESQCRKQQSHLSSIVTPEEQ----EFV 2164

Qy         66 ENLLPSDGDF-WIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEP----SCGSEV 120
              |   :  |: ||||        |    :  : |:||   || ||  ::|    : | :
Db       2165 NN---NAQDYQWIGL--------NDKTIEGDFRWSDGHSLQFENWRPNQPDNFFATGEDC 2213

Qy        121 CVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYS----DEKPAVPSREAEGEETE 175
              |:::|:             :|||  || :  | ||       | | |      |:: :
Db       2214 VVMIWHEKG----------EWNDVPCNYQLPFTCKKGTVACGEPPVVEHARIFGQKKD 2261


RESULT 13
A39086
```

aggrecan precursor, cartilage long splice form [validated] - human
N;Alternate names: chondroitin sulfate proteoglycan 1; large aggregating proteoglycan; p
N;Contains: aggrecan cartilage short splice form
C;Species: Homo sapiens (man)
C;Date: 10-Sep-1999 #sequence_revision 01-Dec-2000 #text_change 08-Dec-2000
C;Accession: A39086; S50206; A43919; S46659; S66389; S68646; S62786; A34226; B43919; C43
R;Doege, K.J.; Sasaki, M.; Kimura, T.; Yamada, Y.
J. Biol. Chem. 266, 894-902, 1991
A;Title: Complete coding sequence and deduced primary structure of the human cartilage l
A;Reference number: A39086; MUID:91093289; PMID:1985970
A;Accession: A39086
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-2162,2201-2329,'A',2392-2415 <DOE>
A;Cross-references: GB:M55172; NID:g178258; PIDN:AAA62824.1; PID:g178259
R;Glumoff, V.; Savontaus, M.; Vehanen, J.; Vuorio, E.
Biochim. Biophys. Acta 1219, 613-622, 1994
A;Title: Analysis of aggrecan and tenascin gene expression in mouse skeletal tissues by
A;Reference number: S50206; MUID:95035091; PMID:7524681
A;Accession: S50206
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 350-497 <GLU>
A;Cross-references: EMBL:X80278; NID:g516295
A;Note: this translation is not annotated in GenBank entry HSAGGREC, release 113.0
R;Sandy, J.D.; Flannery, C.R.; Neame, P.J.; Lohmander, L.S.
J. Clin. Invest. 89, 1512-1516, 1992
A;Title: The structure of aggrecan fragments in human synovial fluid. Evidence for the i
domain.
A;Reference number: A43919; MUID:92235266; PMID:1569188
A;Accession: A43919
A;Molecule type: protein
A;Residues: 361-370,'X',372-373;393-399,'X',401-407,'X',409 <SAN>
A;Cross-references: PIDN:AAB22079.1; PID:g248844; PIDN:AAB22077.1; PID:g248842; PIDN:AAB
A;Experimental source: synovial fluid
A;Note: sequences modified after extraction from NCBI backbone
R;Barry, F.P.; Neame, P.J.; Sasse, J.; Pearson, D.
Matrix Biol. 14, 323-328, 1994
A;Title: Length variation in the keratan sulfate domain of mammalian aggrecan.
A;Reference number: I46998; MUID:95128522; PMID:7827755
A;Accession: S46659
A;Molecule type: DNA
A;Residues: 764-765,'A',767-846,'V',848-862,'X',864 <BAR>
A;Cross-references: EMBL:S74659; NID:g807127; PIDN:AAC60643.1; PID:g807128
A;Note: the authors translated the codon GAA for residue 803 as Ala and CYT for residue
R;Ilic, M.Z.; Mok, M.T.; Williamson, O.D.; Campbell, M.A.; Hughes, C.E.; Handley, C.J.
Arch. Biochem. Biophys. 322, 22-30, 1995
A;Title: Catabolism of aggrecan by explant cultures of human articular cartilage in the
A;Reference number: S66389; MUID:96004775; PMID:7574678
A;Accession: S66389
A;Status: preliminary
A;Molecule type: protein
A;Residues: 17-23;24,'X',26-27;393-401;402-403 <ILI>
R;Fosang, A.J.; Last, K.; Knaeuper, V.; Murphy, G.; Neame, P.J.
FEBS Lett. 380, 17-20, 1996
A;Title: Degradation of cartilage aggrecan by collagenase-3 (MMP-13).
A;Reference number: S68646; MUID:96181659; PMID:8603731
A;Accession: S68646
A;Status: preliminary
A;Molecule type: protein
A;Residues: 'V',404-405,'XX' <FOS>
R;Dudhia, J.; Davidson, C.M.; Wells, T.M.; Vynios, D.H.; Hardingham, T.E.; Bayliss, M.T.
Biochem. J. 313, 933-940, 1996
A;Title: Age-related changes in the content of the C-terminal region of aggrecan in huma
A;Reference number: S62786; MUID:96190740; PMID:8611178
A;Accession: S62786
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: mRNA
A;Residues: 1778-1927,'A',1929-1963,'V',1965-2162,2201-2415 <DUD>
A;Cross-references: EMBL:X17406; NID:g30248; PIDN:CAA35463.1; PID:g30249
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, January 1990
R;Baldwin, C.T.; Reginato, A.M.; Prockop, D.J.
J. Biol. Chem. 264, 15747-15750, 1989
A;Title: A new epidermal growth factor-like domain in the human core protein for the larg
A;Reference number: A34226; MUID:89380154; PMID:2789216
A;Accession: A34226
A;Molecule type: mRNA
A;Residues: 1936-1963,'V',1965-2069,'A',2071-2415 <BAL>
A;Cross-references: GB:J05062; NID:g181167; PIDN:AAA35726.1; PID:g181168
C;Genetics:
A;Gene: GDB:AGC1; CSPG1; CSPGCP; MSK16
A;Cross-references: GDB:127479; OMIM:155760
A;Map position: 15q26-15q26
C;Superfamily: aggrecan; C-type lectin homology; complement factor H repeat homology; EGF
C;Keywords: alternative splicing; cartilage; chondroitin sulfate proteoglycan; extracellu
F;1-19/Domain: signal sequence #status predicted <SIG>
F;20-2415/Product: aggrecan cartilage long splice form #status predicted <MATL>
F;20-2162,2201-2415/Product: aggrecan cartilage short splice form #status predicted <MATN
F;20-2162,2201-2329,'A',2392-2415/Product: aggrecan short splice form #status predicted <
F;44-135/Domain: immunoglobulin homology <IMM>
F;170-247/Domain: link protein repeat homology <LNK1>
F;268-349/Domain: link protein repeat homology <LNK2>
F;495-572/Domain: link protein repeat homology <LNK3>
F;593-673/Domain: link protein repeat homology <LNK4>
F;677-861/Domain: keratan sulfate attachment #status predicted <KSA>
F;864-1510/Domain: chondroitin sulfate attachment #status predicted <CS1>
F;1511-2162/Domain: chondroitin sulfate attachment #status predicted <CS2>
F;2168-2198/Domain: EGF homology <EGF>
F;2205-2325/Domain: C-type lectin homology <LCH>
F;2332-2388/Domain: complement factor H repeat homology <FHD>
F;126,239,333,387,434,602,657,737,1898/Binding site: carbohydrate (Asn) (covalent) #statu
F;371,376/Binding site: keratan sulfate (Thr) (covalent) #status predicted

Query Match 14.2%; Score 158.5; DB 1; Length 2415;
Best Local Similarity 29.1%; Pred. No. 2e-05;
Matches 52; Conservative 22; Mismatches 62; Indels 43; Gaps 11;

```
Qy          9 QPVCRGG---TQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFI 65
              | ||  |    |  ||:  :| |   |  : :|:  ||     | || : :||    :|:
Db       2202 QEVCEEGWNKYQGHCYR--HFPD---RETWVDAERRCREQQSHLSSIVTPEEQ----EFV 2252

Qy         66 ENLLPSDGDF-WIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEPS---CGSEVC 121
                |   :  |: ||||         |    :  : |:||   || ||  ::|     | |
Db       2253 NN---NAQDYQWIGL--------NDRTIEGDFRWSDGHPMQFENWRPNQPDNFFAAGEDC 2301

Qy        122 VVM-YHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYS----DEKPAVPSREAEGEETE 175
              ||| :|:           :|||  ||    | ||       | | |       |:: :
Db       2302 VVMIWHEKG----------EWNDVPCNYHLPFTCKKGTVACGEPPVVEHARTFGQKKD 2349
```

RESULT 14
B42755
E-selectin precursor - mouse
N;Alternate names: endothelial leukocyte adhesion molecule 1 (ELAM-1)
C;Species: Mus musculus (house mouse)
C;Date: 04-Mar-1993 #sequence_revision 31-Dec-2000 #text_change 02-Aug-2002
C;Accession: S23174; B42755
R;Becker-Andre, M.; van Huijsduijnen, R.H.; Losberger, C.; Whelan, J.; Delamarter, J.F.
Eur. J. Biochem. 206, 401-411, 1992
A;Title: Murine endothelial leukocyte-adhesion molecule 1 is a close structural and funct
A;Reference number: S23174; MUID:92283265; PMID:1375914
A;Accession: S23174
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-612 <BEC>
A;Cross-references: GB:M80778; NID:g193014; PIDN:AAA37547.1; PID:g193015
R;Weller, A.; Isenmann, S.; Vestweber, D.
J. Biol. Chem. 267, 15176-15183, 1992
A;Title: Cloning of the mouse endothelial selectins. Expression of both E- and P-selectin
A;Reference number: A42755; MUID:92340571; PMID:1378846
A;Accession: B42755
A;Status: nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 'MKATAGV',1-389,391-612 <WEL>
A;Cross-references: GB:M87862; NID:g193107
A;Experimental source: endothelial cells

```
A;Note: sequence extracted from NCBI backbone (NCBIP:109470)
A;Note: the sequence in GenBank entry MUSESELEC, release 117.0, (PIDN:AAA37577.1; PID:g1
A;Note: it is uncertain whether the initiator is Met-1 or the AUG codon preceding that
C;Superfamily: P-selectin; C-type lectin homology; complement factor H repeat homology;
C;Keywords: glycoprotein; transmembrane protein
F;1-21/Domain: signal sequence #status predicted <SIG>
F;12-138/Domain: C-type lectin homology <LCH>
F;22-612/Product: P-selectin #status predicted <MAT>
F;143-174/Domain: EGF homology <EGF>
F;180-238/Domain: complement factor H repeat homology <FH1>
F;243-300/Domain: complement factor H repeat homology <FH2>
F;305-363/Domain: complement factor H repeat homology <FH3>
F;368-426/Domain: complement factor H repeat homology <FH4>
F;431-489/Domain: complement factor H repeat homology <FH5>
F;494-548/Domain: complement factor H repeat homology <FH6>
F;25,391,528/Binding site: carbohydrate (Asn) (covalent) #status predicted

  Query Match             13.9%;  Score 154.5;  DB 2;  Length 612;
  Best Local Similarity   27.9%;  Pred. No. 9e-06;
  Matches   38;  Conservative   28;  Mismatches   45;  Indels   25;  Gaps    5;

Qy         25 YFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFWIGLRRREE 84
              |:: :|  : ::||   |:||   ||:|::::|   |     ||  |   :|||:|:
Db         23 YYNASSELMTYDEASAYCQRDYTHLVAIQNKEE---INYLNSNLKHSPSYYWIGIRK--- 76

Qy         85 KQSNSTACQDLYAWTDGS---ISQFRNWYVDEPS--CGSEVCVVMYHQPSAPAGIGGPYM 139
                  ::: |        : :||    ||:    :| || :| | :  :|:
Db         77 -------VNNVWIWVGTGKPLTEEAQNWAPGEPNNKQRNEDCVEIYIQRTKDSGM----- 124

Qy        140 FQWNDDRCNMKNNFIC 155
               |||:||| |   :|
Db        125 --WNDERCNKKKLALC 138


RESULT 15
A28452
proteoglycan core protein precursor, cartilage - rat
N;Alternate names: aggrecan
C;Species: Rattus norvegicus (Norway rat)
C;Date: 30-Jun-1989 #sequence_revision 30-Jun-1989 #text_change 21-Jan-2000
C;Accession: A92623; A23835; A28453; A28095; A28452
R;Doege, K.; Sasaki, M.; Horigan, E.; Hassell, J.R.; Yamada, Y.
J. Biol. Chem. 262, 17757-17767, 1987
A;Title: Complete primary structure of the rat cartilage proteoglycan core protein deduc
A;Reference number: A92623; MUID:88087070; PMID:3693370
A;Accession: A92623
A;Molecule type: mRNA
A;Residues: 1-2124 <DOE>
R;Doege, K.; Sasaki, M.; Horigan, E.; Hassell, J.R.; Yamada, Y.
J. Biol. Chem. 263, 10040a, 1988
A;Reference number: A30069
A;Contents: annotation; revision to residue 698
R;Doege, K.; Fernandez, P.; Hassell, J.R.; Sasaki, M.; Yamada, Y.
J. Biol. Chem. 261, 8108-8111, 1986
A;Title: Partial cDNA sequence encoding a globular domain at the C terminus of the rat c
A;Reference number: A23835; MUID:86250698; PMID:2424893
A;Accession: A23835
A;Molecule type: mRNA
A;Residues: 1856-2124 <DO2>
A;Cross-references: GB:M13518; NID:g206104; PIDN:AAA41836.1; PID:g206105
R;Neame, P.J.; Christner, J.E.; Baker, J.R.
J. Biol. Chem. 262, 17768-17778, 1987
A;Title: Cartilage proteoglycan aggregates. The link protein and proteoglycan amino-term
A;Reference number: A28453; MUID:88087071; PMID:3693371
A;Accession: A28453
A;Molecule type: protein
A;Residues: 20-37,'W',39-60,'E',62-64,'X',66-69;70-83;84,89-148,'L',150-238,'S',240,'A',
C;Superfamily: aggrecan; C-type lectin homology; complement factor H repeat homology; EG
C;Keywords: glycoprotein
F;1-19/Domain: signal sequence #status predicted <SIG>
F;20-2124/Product: proteoglycan core protein #status predicted <MAT>
F;44-135/Domain: immunoglobulin homology <IMM>
F;170-247/Domain: link protein repeat homology <LNK1>
F;268-349/Domain: link protein repeat homology <LNK2>
F;504-581/Domain: link protein repeat homology <LNK3>
F;602-683/Domain: link protein repeat homology <LNK4>
F;1914-2034/Domain: C-type lectin homology <LCH>
F;2041-2097/Domain: complement factor H repeat homology <FHD>
F;126,239,333,387,611,667,1842/Binding site: carbohydrate (Asn) (covalent) #status predic

  Query Match             13.8%;  Score 153.5;  DB 2;  Length 2124;
  Best Local Similarity   26.3%;  Pred. No. 4.7e-05;
  Matches   47;  Conservative   27;  Mismatches   62;  Indels   43;  Gaps    9;

Qy          9 QPVCRGG---TQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFI 65
              |  |  |    |  ||:  :| |    |  : :|:  ||      | || : :||: : |
Db       1911 QEQCEEGWTKFQGHCYR--HFPD---RETWVDAERRCREQQSHLSSIVTPEEQEFVNKNA 1965

Qy         66 ENLLPSDGDF-WIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEP----SCGSEV 120
              :       |: ||||        |    :  : |:||    ||  |  ::|     : | :
Db       1966 Q-------DYQWIGL--------NDRTIEGDFRWSDGHSLQFEKWRPNQPDNFFATGEDC 2010

Qy        121 CVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYS----DEKPAVPSREAEGEETE 175
               |:::|:              :|||  || :  | ||       | |||      |:: :
Db       2011 VVMIWHERG-----------EWNDVPCNYQLPFTCKKGTVACGEPPAVEHARTLGQKKD 2058


RESULT 16
A54423
brevican precursor - bovine
C;Species: Bos primigenius taurus (cattle)
C;Date: 09-Sep-1994 #sequence_revision 09-Sep-1994 #text_change 21-Jan-2000
C;Accession: A54423; S41914
R;Yamada, H.; Watanabe, K.; Shimonaka, M.; Yamaguchi, Y.
J. Biol. Chem. 269, 10119-10126, 1994
A;Title: Molecular cloning of brevican, a novel brain proteoglycan of the aggrecan/versi
A;Reference number: A54423; MUID:94193597; PMID:8144512
A;Accession: A54423
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-912 <YAM>
A;Cross-references: GB:X75887; NID:g452820; PIDN:CAA53481.1; PID:g452821
C;Superfamily: aggrecan; C-type lectin homology; complement factor H repeat homology; EG
F;50-139/Domain: immunoglobulin homology <IMM>
F;174-251/Domain: link protein repeat homology <LNK1>
F;272-353/Domain: link protein repeat homology <LNK2>
F;651-682/Domain: EGF homology <EGF>
F;689-809/Domain: C-type lectin homology <LCH>
F;816-872/Domain: complement factor H repeat homology <FHD>

  Query Match             13.6%;  Score 152;  DB 2;  Length 912;
  Best Local Similarity   28.8%;  Pred. No. 2.4e-05;
  Matches   47;  Conservative   17;  Mismatches   61;  Indels   38;  Gaps    8;

Qy         17 QRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFW 76
              |  |||    | ::|| ::|||:  ||  |  | || : :||  |                |
Db        697 QGACYK----HFSARR-SWEEAENKCRMYGAHLASISTPEEQDFINNRYREYQ------W 745

Qy         77 IGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEPS---CGSEVCVVM-YHQPSAPA 132
              |||        |    :  : |:||    : ||   :|        | |||| :|
Db        746 IGL--------NDRTIEGDFLWSDGVPLLYENWNPGQPDSYFLSGENCVVMVWHDQG--- 794

Qy        133 GIGGPYMFQWNDDRCNMKNNFICKYS----DEKPAVPSREAEG 171
                  ||:|  ||   :: ||         | :|  |  |
Db        795 --------QWSDVPCNYHLSYTCKMGLVSCGPPPELPLAEVFG 829


RESULT 17
T24425
hypothetical protein T04A8.3 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C;Accession: T24425
R;Palmer, S.
submitted to the EMBL Data Library, August 1994
A;Reference number: Z19889
```

A;Accession: T24425
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-459 <WIL>
A;Cross-references: EMBL:Z35663; PIDN:CAA84726.1; GSPDB:GN00021; CESP:T04A8.3
A;Experimental source: clone T04A8
C;Genetics:
A;Gene: CESP:T04A8.3
A;Map position: 3
A;Introns: 25/2; 43/2; 130/1; 272/1; 326/1

Query Match 13.4%; Score 149.5; DB 2; Length 459;
Best Local Similarity 28.5%; Pred. No. 1.8e-05;
Matches 49; Conservative 23; Mismatches 52; Indels 48; Gaps 9;

```
Qy         20 CYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDF--WI 77
              |:|  | || |  :| ::|:: |    |  | |  : :| ||::: |  : |:| :   |:
Db        287 CHK--YVHDMS--MNNDDAEKKCNDMGAHLSSFTTYEELKLLDEMILEVYPNDNNIAVWL 342

Qy         78 GLRRREEKQSNSTACQDL----------------------YAWTDGSISQ---------FR 107
              | :||||       | ||                      : | :| ::|         |
Db        343 GAKRREE-------CGDLSKNFTGGYSKDIHDPCARSRVFEWQNG-VAQNPPIFVGDGFD 394

Qy        108 NW---YVDEPSCGSEVCVV-MYHQPSAPAGIGGPYMFQWNDDRCNMKNNFIC 155
               |    |    |  || |:| |    |   |   |   | ||  || :  |:|
Db        395 YWAEKYEPNHSTDSERCLVQMSGSMSVWYGDNKPRNMQINDIYCNYEFKFLC 446
```

RESULT 18
LNHUER
IgE Fc receptor II, low-affinity [validated] - human
N;Alternate names: Blast-2; CD23; Fc-epsilon-RII; lymphocyte IgE receptor
N;Contains: IgE Fc receptor II, splice form a; IgE Fc receptor II, splice form a'; IgE F
C;Species: Homo sapiens (man)
C;Date: 31-Mar-1988 #sequence_revision 31-Mar-1988 #text_change 15-Sep-2000
C;Accession: A26067; S03279; S39442; S39443; A26164; A26589; A31924; JL0132; S29107
R;Kikutani, H.; Inui, S.; Sato, R.; Barsumian, E.L.; Owaki, H.; Yamasaki, K.; Kaisho, T.
Cell 47, 657-665, 1986
A;Title: Molecular structure of human lymphocyte receptor for immunoglobulin E.
A;Reference number: A26067; MUID:87051737; PMID:2877743
A;Accession: A26067
A;Molecule type: mRNA
A;Residues: 1-321 <KIK>
A;Cross-references: GB:M14766; NID:g182449; PIDN:AAA52435.1; PID:g182450
A;Experimental source: EBV-transformed B lymphoblastoid cells RPMI-8866
R;Suter, U.; Bastos, R.; Hofstetter, H.
Nucleic Acids Res. 15, 7295-7308, 1987
A;Title: Molecular structure of the gene and the 5'-flanking region of the human lymphoc
A;Reference number: S03279; MUID:88015596; PMID:2958779
A;Accession: S03279
A;Status: nucleic acid sequence not shown; not compared with conceptual translation
A;Molecule type: DNA
A;Residues: 157-284 <SUT>
A;Cross-references: GB:X06049; NID:g31316
A;Note: all exon sequences were determined but the complete sequence is not shown
R;Matsui, M.; Nunez, R.; Sachi, Y.; Lynch, R.G.; Yodoi, J.
FEBS Lett. 335, 51-56, 1993
A;Title: Alternative transcripts of the human CD23/Fc-epsilon-RII. A possible novel mech
A;Reference number: S39442; MUID:94063078; PMID:8243664
A;Accession: S39442
A;Molecule type: DNA
A;Residues: 1-7,'D',47-50 <MAS1>
A;Experimental source: splice form a'
A;Accession: S39443
A;Molecule type: DNA
A;Residues: 'MNPPSQD',47-50 <MAS2>
A;Experimental source: splice form b'
R;Ludin, C.; Hofstetter, H.; Sarfati, M.; Levy, C.A.; Suter, U.; Alaimo, D.; Kilchherr,
EMBO J. 6, 109-114, 1987
A;Title: Cloning and expression of the cDNA coding for a human lymphocyte IgE receptor.
A;Reference number: A26164; MUID:87218454; PMID:3034567
A;Accession: A26164
A;Molecule type: mRNA
A;Residues: 1-268,'T',270-321 <LUD>
A;Cross-references: GB:X04772; NID:g34002; PIDN:CAA28465.1; PID:g34003
A;Note: the codon given for 269-Asn (ACC) is inconsistent with the authors' translation
A;Note: part of this sequence, including the amino end of soluble forms of the protein, v
R;Ikuta, K.; Takami, M.; Kim, C.W.; Honjo, T.; Miyoshi, T.; Tagaya, Y.; Kawabe, T.; Yodoi
Proc. Natl. Acad. Sci. U.S.A. 84, 819-823, 1987
A;Title: Human lymphocyte Fc receptor for IgE: sequence homology of its cloned cDNA with
A;Reference number: A26589; MUID:87118255; PMID:2949326
A;Accession: A26589
A;Molecule type: mRNA
A;Residues: 1-321 <IKU>
A;Cross-references: GB:M15059; NID:g182447; PIDN:AAA52434.1; PID:g182448
A;Note: part of this sequence, including the amino end of soluble forms of the protein, v
R;Yokota, A.; Kikutani, H.; Tanaka, T.; Sato, R.; Barsumian, E.L.; Suemura, M.; Kishimoto
Cell 55, 611-618, 1988
A;Title: Two species of human Fc-epsilon receptor II (Fc-epsilon-RII/CD23): tissue-specif
A;Reference number: A31924; MUID:89028672; PMID:2972386
A;Accession: A31924
A;Molecule type: mRNA
A;Residues: 'MNPPSQ',8-14 <YOK>
A;Cross-references: GB:M23562; NID:g182444
A;Experimental source: splice form IIb
R;Letellier, M.; Sarfati, M.; Delespesse, G.
Mol. Immunol. 26, 1105-1112, 1989
A;Title: Mechanisms of formation of IgE-binding factors (soluble CD23)-I. Fc epsilon R II
A;Reference number: JL0132; MUID:90220658; PMID:2534424
A;Accession: JL0132
A;Molecule type: protein
A;Residues: 1-321 <LET>
A;Experimental source: lymphoblastoid B cell line
R;Rose, K.; Turcatti, G.; Graber, P.; Pochon, S.; Regamey, P.O.; Jansen, K.U.; Magnenat,
Biochem. J. 286, 819-824, 1992
A;Title: Partial characterization of natural and recombinant human soluble CD23.
A;Reference number: S29107; MUID:93038513; PMID:1417742
A;Accession: S29107
A;Molecule type: protein
A;Residues: 152-166;173-179;189-212;230-263;268-306 <ROS>
R;Padlan, E.A.; Helm, B.A.
submitted to the Brookhaven Protein Data Bank, June 1993
A;Reference number: A51791; PDB:1HLI
A;Contents: annotation; conformation by theoretical model, residues 173-285
R;Bajorath, J.
submitted to the Brookhaven Protein Data Bank, November 1995
A;Reference number: A65963; PDB:1KJE
A;Contents: annotation; conformation by theoretical model, residues 173-285
C;Comment: The sequence of the splice form a is shown.
C;Comment: This receptor for the Fc portion of IgE is expressed in various hematopoietic
ells.
C;Comment: Splice form a is expressed constitutively in B-cells; b is expressed in other
C;Comment: Soluble IgE-binding factors are produced by proteolytic cleavage of IgE Fc rec
C;Genetics:
A;Gene: GDB:FCER2; FCE2
A;Cross-references: GDB:118888; OMIM:151445
A;Map position: 19p13.3-19p13.3
A;Introns: 8/1; 46/1; 64/1; 85/1; 106/1; 127/1; 157/1; 207/3; 243/2
C;Superfamily: IgE receptor II; C-type lectin homology
C;Keywords: alternative splicing; B-cell; glycoprotein; immunoglobulin receptor; macropha
F;'MNPPSQD',47-321/Product: IgE Fc receptor II, splice form b' #status predicted <SFBI>
F;'MNPPSQ',8-321/Product: IgE Fc receptor II, splice form b #status predicted <SFB>
F;1-321/Product: IgE Fc receptor II, splice form a #status predicted <SFA>
F;1-23/Domain: intracellular #status predicted <INT>
F;1-7,'D',47-321/Product: IgE Fc receptor II, splice form a' #status predicted <SFAI>
F;14-20/Region: stop-transfer sequence
F;22-45/Domain: transmembrane #status predicted <TMM>
F;46-321/Domain: extracellular #status predicted <EXT>
F;64-84/Region: 21-residue repeat
F;81-321/Product: soluble IgE-binding factor (37K) #status predicted <IGE>
F;85-105/Region: 21-residue repeat
F;102-321/Product: soluble IgE-binding factor (33K) #status predicted <IGB>
F;106-126/Region: 21-residue repeat
F;125-321/Product: soluble IgE-binding factor (29K) #status predicted <IGI>
F;148-321/Product: soluble IgE-binding factor (25-27K), long form #status experimental <I
F;150-321/Product: soluble IgE-binding factor (25-27K), short form #status experimental <

```
F;163-282/Domain: C-type lectin homology <LCH>
F;63/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;147-148/Cleavage site: Lys-Leu (unidentified proteinase) #status experimental
F;149-150/Cleavage site: Arg-Met (unidentified proteinase) #status experimental
F;191-282,259-273/Disulfide bonds: #status experimental

  Query Match             13.4%;  Score 149;  DB 1;  Length 321;
  Best Local Similarity   30.2%;  Pred. No. 1.3e-05;
  Matches   55;  Conservative   20;  Mismatches   57;  Indels   50;  Gaps   11;

Qy         17 QRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFW 76
              || ||   ||   :::  :  |: ||    |||||| | :||  : |       |    |
Db        171 QRKCY---YFGKGTKQ--WVHARYACDDMEGQLVSIHSPEEQDFLTKH-----ASHTGSW 220

Qy         77 IGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEPSCGS--EVCVVMYHQPSAPAGI 134
              ||||  : |          :  |||   : ||   ||:  |  | ||:|       |
Db        221 IGLRNLDLKGE--------FIWVDGSHVDYSNWAPGEPTSRSQGEDCVMM-------RGS 265

Qy        135 GGPYMFQWNDDRCNMK-NNFICKYSDEKPAV---PSREAEGE----------ETELTTPV 180
              |     :|||  |: |   ::|    :: |    |: |   |          :  | ||
Db        266 G-----RWNDAFCDRKLGAWVC----DRLATCTPPASEGSAESMGPDSRPDPDGRLPTPS 316

Qy        181 LP 182
              |
Db        317 AP 318


RESULT 19
T46256
brevican - human (fragment)
N;Alternate names: protein DKFZp761L191.1
C;Species: Homo sapiens (man)
C;Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 02-Sep-2000
C;Accession: T46256
R;Ottenwaelder, B.; Obermaier, B.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
submitted to the Protein Sequence Database, January 2000
A;Reference number: Z23031
A;Accession: T46256
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-330 <AAA>
A;Cross-references: EMBL:AL137504
A;Experimental source: adult amygdala; clone DKFZp761L191
C;Genetics:
A;Note: DKFZp761L191.1
C;Superfamily: aggrecan; C-type lectin homology; complement factor H repeat homology; EG

  Query Match             13.3%;  Score 148.5;  DB 2;  Length 330;
  Best Local Similarity   29.5%;  Pred. No. 1.5e-05;
  Matches   46;  Conservative   16;  Mismatches   57;  Indels   37;  Gaps    8;

Qy          8 GQPVCRGG---TQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEKF 64
              |   |  |    |  |||    | ::|| ::|||:  ||  |  | || : :||  |
Db        103 GLRFCNPGWDAFQGACYK----HFSTRR-SWEEAETQCRMYGAHLASISTPEEQDFINNR 157

Qy         65 IENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEPS---CGSEVC 121
                 ||||        |    :  : |:||    : ||   :|       | |
Db        158 YREYQ------WIGL--------NDRTIEGDFLWSDGVPLLYENWNPGQPDSYFLSGENC 203

Qy        122 VVM-YHQPSAPAGIGGPYMFQWNDDRCNMKNNFICK 156
              ||| :|              ||:|  ||   :: ||
Db        204 VVMVWHDQG----------QWSDVPCNYHLSYTCK 228


RESULT 20
A55182
aggrecan precursor - mouse
C;Species: Mus musculus (house mouse)
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 21-Jan-2000
C;Accession: A55182; S55329; S50207; S51355; I78532; I58123
R;Walcz, E.; Deak, F.; Erhardt, P.; Coulter, S.N.; Fueloep, C.; Horvath, P.; Doege, K.J.
Genomics 22, 364-371, 1994
A;Title: Complete coding sequence, deduced primary structure, chromosomal localization,
A;Reference number: A55182; MUID:95104847; PMID:7806222
A;Accession: A55182
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: mRNA
A;Residues: 1-2132 <WAL>
A;Cross-references: GB:L07049; NID:g678541; PIDN:AAC37670.1; PID:g191772
R;Watanabe, H.; Gao, L.; Sugiyama, S.; Doege, K.; Kimata, K.; Yamada, Y.
Biochem. J. 308, 433-440, 1995
A;Title: Mouse aggrecan, a large cartilage proteoglycan: protein sequence, gene structur
A;Reference number: S55329; MUID:95289972; PMID:7772024
A;Accession: S55329
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-567,'G',569-1728,'I',1730-2132 <WAT1>
A;Cross-references: GB:U22901; NID:g886014
R;Glumoff, V.; Savontaus, M.; Vehanen, J.; Vuorio, E.
Biochim. Biophys. Acta 1219, 613-622, 1994
A;Title: Analysis of aggrecan and tenascin gene expression in mouse skeletal tissues by
A;Reference number: S50206; MUID:95035091; PMID:7524681
A;Accession: S50207
A;Molecule type: mRNA
A;Residues: 350-481,'R',483-506 <GLU1>
A;Cross-references: EMBL:X80279; NID:g673432
R;Glumoff, V.
submitted to the EMBL Data Library, July 1994
A;Reference number: S51355
A;Accession: S51355
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 350-383,'CPVMSQRERPWAA' <GLU2>
A;Cross-references: EMBL:X80279
R;Watanabe, H.; Kimata, K.; Line, S.; Strong, D.; Gao, L.Y.; Kozak, C.A.; Yamada, Y.
Nature Genet. 7, 154-157, 1994
A;Title: Mouse cartilage matrix deficiency (cmd) caused by a 7 bp deletion in the aggrec
A;Reference number: I58123; MUID:95004579; PMID:7920633
A;Accession: I78532
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 211-326 <WAT2>
A;Cross-references: GB:S73722; NID:g765215; PIDN:AAB32160.1; PID:g765216
A;Accession: I58123
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 211-240,'MCTASLRRWRVRSFMRHPQRNSPSRRQPTS','AGGWGHAWPPQASSTWPGRAVWTCAALAGW','R
A;Cross-references: GB:S73720; NID:g765211; PIDN:AAB32159.1; PID:g765212
C;Genetics:
A;Map position: 7
A;Introns: 253/1
C;Superfamily: aggrecan; C-type lectin homology; complement factor H repeat homology; EG
C;Keywords: cartilage; extracellular matrix
F;1-19/Domain: signal sequence #status predicted <SIG>
F;44-135/Domain: immunoglobulin homology <IMM>
F;170-247/Domain: link protein repeat homology <LNK1>
F;268-349/Domain: link protein repeat homology <LNK2>
F;504-581/Domain: link protein repeat homology <LNK3>
F;602-683/Domain: link protein repeat homology <LNK4>
F;1922-2042/Domain: C-type lectin homology <LCH>
F;2049-2105/Domain: complement factor H repeat homology <FHD>

  Query Match             13.3%;  Score 148.5;  DB 1;  Length 2132;
  Best Local Similarity   24.7%;  Pred. No. 0.00013;
  Matches   46;  Conservative   30;  Mismatches   67;  Indels   43;  Gaps    9;

Qy          2 TGRLLSGQPVCRGG---TQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQ 58
              |   :: |  | |     |  ||:  :| |   |  : :|:  ||      | || : :||
Db       1912 TESTVADQEQCEEGWTKFQGHCYR--HFPD---RETWVDAERRCREQQSHLSSIVTPEEQ 1966

Qy         59 KLIEKFIENLLPSDGDF-WIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEP--- 114
              : : |  :        |: ||||         |    :  : |:||    ||   |  ::|
Db       1967 EFVNKNAQ-------DYQWIGL--------NDRTIEGDFRWSDGHSLQFEKWRPNQPDNF 2011

Qy        115 -SCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYS----DEKPAVPSREA 169
               : | :  |:::|:              :|||  || :  | ||       : | |
```

```
Db       2012 FATGEDCVVMIWHERG----------EWNDVPCNYQLPFTCKKGTVACGDPPVVEHART 2060

Qy        170 EGEETE 175
              |:: :
Db       2061 LGQKKD 2066


RESULT 21
E89130
protein F52E1.2 [imported] - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 10-May-2001
C;Accession: E89130
R;anonymous, The C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998
A;Title: Genome sequence of the nematode C. elegans: a platform for investigating biolog
A;Reference number: A75000; MUID:99069613; PMID:9851916
A;Note: see websites genome.wustl.edu/gsc/C_elegans/ and www_sanger.ac.uk/Projects/C_ele
A;Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and
A;Accession: E89130
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-253 <STO>
A;Cross-references: GB:chr_V; PIDN:AAB37037.1; PID:g1086809; GSPDB:GN00023; CESP:F52E1.2
C;Genetics:
A;Gene: F52E1.2
A;Map position: 5

  Query Match             13.3%;  Score 148;  DB 2;  Length 253;
  Best Local Similarity   29.2%;  Pred. No. 1.2e-05;
  Matches   49;  Conservative   24;  Mismatches   57;  Indels   38;  Gaps   10;

Qy         12 CRGG-------TQRPCYKV-IYFHDTSRRLNFEE---------AKEACRRDGGQLVSIES 54
              | ||           |||| |:||  | ::  ::          |  ||    | :||:|:|
Db         96 CPGGCPTGWQYLNSKCYKVLIHFHCLSLKVQLQKFDAAVTYAGATSACAAQGAELVTIDS 155

Qy         55 EDEQKLIEKFIEN--LLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVD 112
               ||   : |  :   |:    : ||||:       | | | |     | ||| : : ||
Db        156 FDENDALRKAFDTNALVDETKETWIGLK------SLSGAWQ----WADGSSATYTNWAPS 205

Qy        113 EPSCGSEVCVVMYHQPSAPAGIGGPYMFQ---WNDDRCNMKN-NFICK 156
              :||  : :|| |     : |     | :|   |    |   : ::||:
Db        206 QPS-SNGLCVQMITDSLSNA----TYKYQRGGWKTYGCGKTSASYICE 248


RESULT 22
S57653
brevican precursor - mouse
C;Species: Mus musculus (house mouse)
C;Date: 19-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 13-Aug-1999
C;Accession: S57653
R;Rauch, U.; Forsberg, N.; Kulbe, G.; Arnold-Ammer, I.; Faessler, R.
submitted to the EMBL Data Library, May 1995
A;Description: Cloning and sequence of mouse neurocan and brevican and their different e
A;Reference number: S57653
A;Accession: S57653
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-883 <RAU>
A;Cross-references: EMBL:X87096; NID:g886889; PIDN:CAA60575.1; PID:g886890
C;Superfamily: aggrecan; C-type lectin homology; complement factor H repeat homology; EG
F;49-138/Domain: immunoglobulin homology <IMM>
F;173-250/Domain: link protein repeat homology <LNK1>
F;271-352/Domain: link protein repeat homology <LNK2>
F;626-657/Domain: EGF homology <EGF>
F;664-784/Domain: C-type lectin homology <LCH>
F;791-847/Domain: complement factor H repeat homology <FHD>

  Query Match             13.2%;  Score 147;  DB 2;  Length 883;
  Best Local Similarity   29.7%;  Pred. No. 6.4e-05;
  Matches   43;  Conservative   21;  Mismatches   45;  Indels   36;  Gaps    8;

Qy         17 QRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLI-EKFIENLLPSDGDF 75
                |  |||    | ::|| ::|||:  ||  |  | || : :||  : :::  |
Db        672 QGACYK----HFSTRR-SWEEAESQCRALGAHLTSICTPEEQDFVNDRYREYQ------- 719

Qy         76 WIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEPS---CGSEVCVVM-YHQPSAP 131
              ||||         |    :  : |:||:   : ||   :|       | |||| :|
Db        720 WIGL-------NDRTIEGDFLWSDGAPLLYENWNPGQPDSYFLSGENCVVMVWHDQG-- 769

Qy        132 AGIGGPYMFQWNDDRCNMKNNFICK 156
                 ||:|  ||   :: ||
Db        770 ---------QWSDVPCNYHLSYTCK 785


RESULT 23
LNRC1
lectin BRA3-1 precursor - barnacle (Megabalanus rosa)
C;Species: Megabalanus rosa
C;Date: 24-Feb-1994 #sequence_revision 09-Sep-1994 #text_change 16-Jul-1999
C;Accession: JC1503; A26094
R;Takamatsu, N.; Takeda, T.; Kojima, M.; Heishi, M.; Muramoto, K.; Kamiya, H.; Shiba, T.
Gene 128, 251-255, 1993
A;Title: Acorn barnacle Megabalanus rosa lectin (BRA-3): cDNA cloning, gene structure and
A;Reference number: JC1503; MUID:93292994; PMID:8514190
A;Accession: JC1503
A;Molecule type: DNA; mRNA
A;Residues: 1-162 <TAK>
A;Cross-references: DDBJ:D13299
R;Muramoto, K.; Kamiya, H.
Biochim. Biophys. Acta 874, 285-295, 1986
A;Title: The amino-acid sequence of a lectin of the acorn barnacle Megabalanus rosa.
A;Reference number: A26094
A;Accession: A26094
A;Molecule type: protein
A;Residues: 25-145,'K',147-162 <MUR>
A;Note: 146-Arg was also found
C;Comment: This galactose-binding lectin is isolated from the coelomic fluid.
C;Comment: This protein plays important roles in defense mechanisms and in development ar
C;Comment: The molecule is a tetramer of identical chains.
C;Genetics:
A;Introns: 22/1; 47/2; 86/3
C;Superfamily: tetranectin; C-type lectin homology
C;Keywords: hemolymph; homotetramer; lectin
F;1-24/Domain: signal sequence #status predicted <SIG>
F;25-162/Product: lectin BRA3-1 #status experimental <MAT>
F;26-150/Domain: C-type lectin homology <LCH>
F;26-39,56-150,125-142/Disulfide bonds: #status experimental
F;157/Disulfide bonds: interchain (to 160) #status experimental
F;160/Disulfide bonds: interchain (to 157) #status experimental

  Query Match             13.1%;  Score 146.5;  DB 1;  Length 162;
  Best Local Similarity   25.0%;  Pred. No. 9.8e-06;
  Matches   40;  Conservative   26;  Mismatches   59;  Indels   35;  Gaps    6;

Qy          5 LLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACR--RDGGQLVSIESEDEQKLIE 62
              : :|:  | |       |    :  :: :: : :|: ||:    |  | :|:|: |   |
Db         19 ITTGECTCPGNLDWQEYDGHCYWASTYQVRWNDAQLACQTVHPGAYLATIQSQLENAFIS 78

Qy         63 KFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEPS------C 116
              : : |        ||||       |    :  | |::|  :  |  |  : |:      |
Db         79 ETVSN-----NRLWIGL--------NDIDLEGHYVWSNGEATDFTYWSSNNPNNWENQDC 125

Qy        117 GSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICK 156
              |    || |   :          ||:|| ||   ||:||
Db        126 G----VVNYDTVTG----------QWDDDDCNKNRNFLCK 151


RESULT 24
I50421
aggrecan precursor - chicken
N;Alternate names: cartilage chondroitin sulfate proteoglycan core protein
C;Species: Gallus gallus (chicken)
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 21-Jan-2000
C;Accession: I50421; S39796; S27356; A25442; A32002; I50216; A37072; B37072
R;Li, H.; Schwartz, N.B.; Vertel, B.M.
```

J. Biol. Chem. 268, 23504-23511, 1993
A;Title: cDNA cloning of chick cartilage chondroitin sulfate (aggrecan) core protein and
A;Reference number: A48884; MUID:94043149; PMID:8226878
A;Accession: I50421
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-2109 <LIX>
A;Cross-references: GB:L21913; NID:g416133; PIDN:AAB19128.1; PID:g416134
R;Chandrasekaran, L.; Tanzer, M.L.
Biochem. J. 296, 885-887, 1993
A;Reference number: S39796; MUID:94107258; PMID:8280087
A;Contents: annotation; erratum
A;Accession: S39796
A;Molecule type: mRNA
A;Residues: 1-361,'DL',364-600,'R',602-999,'R',1001-1028,'P',1030-1250,'D',1252-1602,'A'
A;Cross-references: GB:M88101
R;Chandrasekaran, L.; Tanzer, M.L.
Biochem. J. 288, 903-910, 1992
A;Title: Molecular cloning of chicken aggrecan. Structural analyses.
A;Reference number: S27356; MUID:93111968; PMID:1339285
A;Accession: S27356
A;Molecule type: mRNA
A;Residues: 1-361,'DL',364-600,'R',602-999,'R',1001-1028,'P',1030-1250,'D',1252-1549,'T'
A;Cross-references: EMBL:M88101
R;Sai, S.; Tanaka, T.; Kosher, R.A.; Tanzer, M.L.
Proc. Natl. Acad. Sci. U.S.A. 83, 5081-5085, 1986
A;Title: Cloning and sequence analysis of a partial cDNA for chicken cartilage proteogly
A;Reference number: A25442; MUID:86259736; PMID:3460082
A;Accession: A25442
A;Molecule type: mRNA
A;Residues: 1693-1795,'G',1797-1855,1894-2109 <SAI>
A;Cross-references: GB:M13993; NID:g211654; PIDN:AAA48720.1; PID:g211655
A;Experimental source: sternal cartilage
R;Tanaka, T.; Har-El, R.; Tanzer, M.L.
J. Biol. Chem. 263, 15831-15835, 1988
A;Title: Partial structure of the gene for chicken cartilage proteoglycan core protein.
A;Reference number: A32002; MUID:89008500; PMID:3170613
A;Accession: A32002
A;Molecule type: DNA
A;Residues: 1893-1987,'S',1989-2022 <TAN>
A;Note: the authors translated the codon TCC for residue 1787 as Phe
R;Krueger, R.C.
J. Biol. Chem. 265, 12088-12097, 1990
A;Title: Chick cartilage chondroitin sulfate proteoglycan core protein: II. Nucleotide s
A;Reference number: I50216; MUID:90307744; PMID:1694853
A;Accession: I50216
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 'PA',1044-1559 <KRU>
A;Cross-references: GB:M38187; NID:g211685; PIDN:AAA48731.1; PID:g555441
R;Krueger Jr., R.C.; Fields, T.A.; Hildreth IV, J.; Schwartz, N.B.
J. Biol. Chem. 265, 12075-12087, 1990
A;Title: Chick cartilage chondroitin sulfate proteoglycan core protein. I. Generation an
A;Reference number: A37072; MUID:90307743; PMID:2365711
A;Accession: A37072
A;Molecule type: protein
A;Residues: 998-1015,'X',1017-1019,'X',1021-1023 <KR2>
A;Note: amino end of 86K core peptide CS-A
A;Accession: B37072
A;Molecule type: protein
A;Residues: 1247-1250,'D',1252-1272,'X',1274-1275 <KR3>
A;Note: amino end of 75K core peptide CS-B
C;Superfamily: aggrecan; C-type lectin homology; complement factor H repeat homology; EG
C;Keywords: alternative splicing
F;1-20/Domain: signal sequence #status predicted <SIG>
F;21-2109/Product: aggrecan #status predicted <MAT>
F;44-131/Domain: immunoglobulin homology <IMM>
F;166-243/Domain: link protein repeat homology <LNK1>
F;264-346/Domain: link protein repeat homology <LNK2>
F;537-614/Domain: link protein repeat homology <LNK3>
F;635-716/Domain: link protein repeat homology <LNK4>
F;1859-1890/Domain: EGF homology <EGF>
F;1897-2017/Domain: C-type lectin homology <LCH>
F;2024-2080/Domain: complement factor H repeat homology <FHD>

Query Match 13.1%; Score 146; DB 1; Length 2109;
Best Local Similarity 25.0%; Pred. No. 0.00022;
Matches 42; Conservative 28; Mismatches 58; Indels 40; Gaps 8;

```
Qy         17 QRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDF- 75
              |  ||:  :| :   |  : :|:  ||     | || : :||: :    :       |:
Db       1905 QGHCYR--HFEE---RETWMDAESRCREHQAHLSSIITPEEQEFVNSHAQ-------DYQ 1952

Qy         76 WIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEPS----CGSEVCVVMYHQPSAP 131
              ||||  |        | :: : |:||   || ||  ::|     | :  |:::|:
Db       1953 WIGLSDR--------AVENDFRWSDGHSLQFENWRPNQPDNFFFAGEDCVVMIWHEQG-- 2002

Qy        132 AGIGGPYMFQWNDDRCNMKNNFICKYS----DEKPAVPSREAEGEETE 175
                 :|||  ||    | ||       : | | :    | : :
Db       2003 ---------EWNDVPCNYHLPFTCKKGTVACGDPPVVENARTFGRKKD 2041
```

RESULT 25
JC7595
scavenger receptor with C-type lectin type I - human
C;Species: Homo sapiens (man)
C;Date: 30-Jun-2001 #sequence_revision 30-Jun-2001 #text_change 30-Jun-2001
C;Accession: JC7595
R;Nakamura, K.; Funakoshi, H.; Miyamoto, K.; Tokunaga, F.; Nakamura, T.
Biochem. Biophys. Res. Commun. 280, 1028-1035, 2001
A;Title: Molecular cloning and functional characterization of a human scavenger receptor
A;Reference number: JC7595; MUID:21092718; PMID:11162630
A;Contents: Placenta
A;Accession: JC7595
A;Molecule type: mRNA
A;Residues: 1-742 <NAK>
A;Cross-references: DDBJ:AB038518
C;Comment: This receptor, a member of the scavenger receptor family, belonging to the ty
important role in host defense. It forms a timer and plays a role in recognizing infect
C;Genetics:
A;Gene: srcl-I
A;Map position: 18p11.32
C;Keywords: coiled coil; glycoprotein; transmembrane protein
F;1-39/Domain: cytosolic (amino-terminus) #status predicted <CYT>
F;16-19/Region: internalization signal YKRF
F;40-56/Domain: transmembrane #status predicted <TMM>
F;57-112/Domain: extracellular #status predicted <EXT>
F;113-335/Domain: coiled coil #status predicted <COC>
F;369-384/Region: serine/threonine-rich #status predicted
F;443-589/Domain: collagen-like #status predicted <COL>
F;607-732/Domain: C-type lectin/carbohydrate recognition domain #status predicted <CRD>

Query Match 13.0%; Score 145.5; DB 2; Length 742;
Best Local Similarity 28.7%; Pred. No. 7.1e-05;
Matches 39; Conservative 17; Mismatches 61; Indels 19; Gaps 5;

```
Qy         25 YFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFWIGLRRREE 84
              |:    : : ||:||  |       ||  | : :||: |:|     :      ||||   |
Db       619 YYFSVEKEI-FEDAKLFCEDKSSHLVFINTREEQQWIKK----QMVGRESHWIGLTDSER 673

Qy         85 KQSNSTACQDLYAWTDGSISQFRNWYVDEPSCGSEVCVVMYHQPSAPAGIGGPYMFQWND 144
              :          : | ||:   ::||   :|           | |      |  |  ||||
Db       674 ENE--------WKWLDGTSPDYKNWKAGQPDNWGH-----GHGPGEDCA-GLIYAGQWND 719

Qy        145 DRCNMKNNFICKYSDE 160
               :|   |||||:   |
Db       720 FQCEDVNNFICEKDRE 735
```

RESULT 26
S49126
brevican precursor - rat
N;Alternate names: aggrecan-like protein
C;Species: Rattus norvegicus (Norway rat)
C;Date: 01-Feb-1995 #sequence_revision 12-May-1995 #text_change 26-May-2000
C;Accession: S49126; I55457

```
R;Seidenbecher, C.I.; Langnase, K.; Wax, H.; Seidel, B.; Garner, C.C.; Gundelfinger, E.D
submitted to the EMBL Data Library, June 1994
A;Description: Molecular cloning of a new member of the aggrecan/versican family of prot
A;Reference number: S49126
A;Accession: S49126
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-883 <SEI>
A;Cross-references: EMBL:X79881; NID:g509396; PIDN:CAA56255.1; PID:g509397
R;Seidenbecher, I.C.; Richter, K.; Rauch, U.; Fassler, R.; Garner, C.C.; Gundelfinger, E
J. Biol. Chem. 270, 27206-27212, 1995
A;Title: Brevican, a Chondroitin Sulfate Proteoglycan of Rat Brain, Occurs as Secreted a
A;Reference number: I55457; MUID:96070828; PMID:7592978
A;Accession: I55457
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-883 <RES>
A;Cross-references: EMBL:X79881; NID:g509396; PIDN:CAA56255.1; PID:g509397
C;Comment: For an alternative splice form, see PIR:A53908.
C;Superfamily: aggrecan; C-type lectin homology; complement factor H repeat homology; EG
C;Keywords: alternative splicing
F;49-138/Domain: immunoglobulin homology <IMM>
F;173-250/Domain: link protein repeat homology <LNK1>
F;271-352/Domain: link protein repeat homology <LNK2>
F;626-657/Domain: EGF homology <EGF>
F;664-784/Domain: C-type lectin homology <LCH>
F;791-847/Domain: complement factor H repeat homology <FHD>

  Query Match             13.0%;  Score 145;  DB 2;  Length 883;
  Best Local Similarity   29.7%;  Pred. No. 9.6e-05;
  Matches   43;  Conservative   20;  Mismatches   46;  Indels   36;  Gaps    8;

Qy         17 QRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLI-EKFIENLLPSDGDF 75
              |  |||     | ::|| ::|||:  ||  |  | || : :||  : ::: |
Db        672 QGACYK----HFSTRR-SWEEAESQCRALGAHLTSICTPEEQDFVNDRYREYQ------- 719

Qy         76 WIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEPS---CGSEVCVVM-YHQPSAP 131
              ||||        |    :  : |:||    : ||   :|       | |||| :|
Db        720 WIGL--------NDRTIEGDFLWSDGPPLLYENWNPGQPDSYFLSGENCVVMVWHDQG-- 769

Qy        132 AGIGGPYMFQWNDDRCNMKNNFICK 156
                    ||:|  ||   :: ||
Db        770 ---------QWSDVPCNYHLSYTCK 785


RESULT 27
S10548
lectin - barnacle (Megabalanus rosa)
C;Species: Megabalanus rosa
C;Date: 02-Dec-1993 #sequence_revision 26-May-1995 #text_change 26-May-1995
C;Accession: S10548
R;Muramoto, K.; Kamiya, H.
Biochim. Biophys. Acta 1039, 42-51, 1990
A;Title: The amino-acid sequence of multiple lectins of the acorn barnacle Megabalanus r
A;Reference number: S10548; MUID:90283457; PMID:2354200
A;Accession: S10548
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-173 <MUR>
A;Note: the sequence from Fig. 6 is inconsistent with that from Fig. 5 in having additio

  Query Match             13.0%;  Score 144.5;  DB 2;  Length 173;
  Best Local Similarity   28.6%;  Pred. No. 1.6e-05;
  Matches   38;  Conservative   18;  Mismatches   50;  Indels   27;  Gaps    5;

Qy         26 FHDTSRRLNFEEAKEACRR--DGGQLVSIESEDEQKLIEKFIENLLPSDGDFWIGLRRRE 83
              ||    : ::  |   | |     :| ||:: | | ::|        |    ||||
Db         62 FHVPLEKASWMVAHGVCARLDSRARLASIDAAD-QAVVEPL------SSEKMWIGL---- 110

Qy         84 EKQSNSTACQDLYAWTDGSISQFRNWYVDEPSCGSEVCVVMYHQPSAPAGIGGPYMFQWN 143
                 :  :  |   | | | |  ||||  :|   ||:||::                  ||:
Db        111 ----SYDSANDAAVWADDSHSSHRNWYATQPDDESELCVLIKEDQYR----------QWH 156
```

```
Qy        144 DDRCNMKNNFICK 156
                 |  || : ||:|:
Db        157 DYNCNDRYNFVCE 169


RESULT 28
S23936
L-selectin precursor - rat
N;Alternate names: leucocyte cell adhesion molecule 1 (LECAM-1)
C;Species: Rattus norvegicus (Norway rat)
C;Date: 02-Dec-1993 #sequence_revision 01-Sep-1995 #text_change 20-Jun-2000
C;Accession: S23936
R;Watanabe, T.; Song, Y.; Hirayama, Y.; Tamatani, T.; Kuida, K.; Miyasaka, M.
Biochim. Biophys. Acta 1131, 321-324, 1992
A;Title: Sequence and expression of a rat cDNA for LECAM-1.
A;Reference number: S23936; MUID:92329548; PMID:1378303
A;Accession: S23936
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-372 <WAT>
A;Cross-references: GB:D10831; NID:g220801; PIDN:BAA01613.1; PID:g220802
C;Superfamily: L-selectin; C-type lectin homology; complement factor H repeat homology; E
C;Keywords: transmembrane protein
F;29-155/Domain: C-type lectin homology <LCH>
F;160-191/Domain: EGF homology <EGF>
F;197-254/Domain: complement factor H repeat homology <FH1>
F;259-316/Domain: complement factor H repeat homology <FH2>

  Query Match             13.0%;  Score 144.5;  DB 2;  Length 372;
  Best Local Similarity   26.5%;  Pred. No. 3.9e-05;
  Matches   41;  Conservative   35;  Mismatches   52;  Indels   27;  Gaps    8;

Qy         26 FHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGD-FWIGLRRREE 84
              :| : | :|:|  |:: |: :    ||:|::: |   || ::|  || :   :|||:|:
Db         41 YHYSERSMNWENARKFCKHNYTDLVAIQNKRE---IE-YLEKTLPKNPTYYWIGIRK--- 93

Qy         85 KQSNSTACQDLYAW--TDGSIS-QFRNWYVDEPS--CGSEVCVVMYHQPSAPAGIGGPYM 139
                        : |  |: ::: :  ||   ||:     | || :| :    :|
Db         94 -------IGKTWTWVGTNKTLTKEAENWGTGEPNNKKSKEDCVEIYIKRERDSG------ 140

Qy        140 FQWNDDRCNMKNNFICKYSDEKPAVPSREAEGEET 174
                :|||| |: :   :|  :  :|   :|  |  ||
Db        141 -KWNDDACHKRKAALCYTASCQPESCNRHGECVET 174


RESULT 29
A46274
HIV gp120-binding C-type lectin - human
C;Species: Homo sapiens (man)
C;Date: 21-Sep-1993 #sequence_revision 18-Nov-1994 #text_change 01-Dec-1995
C;Accession: A46274
R;Curtis, B.M.; Scharnowske, S.; Watson, A.J.
Proc. Natl. Acad. Sci. U.S.A. 89, 8356-8360, 1992
A;Title: Sequence and expression of a membrane-associated C-type lectin that exhibits CD4
A;Reference number: A46274; MUID:92390446; PMID:1518869
A;Accession: A46274
A;Status: preliminary
A;Molecule type: nucleic acid
A;Residues: 1-404 <CUR>
A;Experimental source: placenta
A;Note: sequence extracted from NCBI backbone (NCBIN:113134, NCBIP:113135)
C;Superfamily: C-type lectin homology
F;256-377/Domain: C-type lectin homology <LCH>

  Query Match             13.0%;  Score 144.5;  DB 2;  Length 404;
  Best Local Similarity   29.7%;  Pred. No. 4.3e-05;
  Matches   47;  Conservative   23;  Mismatches   47;  Indels   41;  Gaps   10;

Qy         25 YFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFWIGLRRREE 84
              ||   |:| |: ::  ||:  | ||| |:| :||  ::        |:   |:||    :
Db        268 YFMSNSQR-NWHDSITACKEVGAQLVVIKSAEEQNFLQLQSSR---SNRFTWMGLSDLNQ 323

Qy         85 KQSNSTACQDLYAWTDGS--ISQFRN-WYVDEP-SCGSEVCVVMYHQPSAPAGIGGPYMF 140
```

```
                    :  : | |||  :  |:  |   || : | | |          :| |
Db        324 --------EGTWQWVDGSPLLPSFKQYWNRGEPNNVGEEDCAEF-------SGNG----- 363

Qy        141 QWNDDRCNMKNNFICKYS------DEK------PAVPS 166
               ||||:||:   :||| |      ||:      || |:
Db        364 -WNDDKCNLAKFWICKKSAASCSRDEEQFLSPAPATPN 400


RESULT 30
JC4329
coagulation factor IX-binding protein A chain - habu
C;Species: Trimeresurus flavoviridis (habu)
C;Date: 06-Dec-1995 #sequence_revision 08-Feb-1996 #text_change 13-Mar-1998
C;Accession: JC4329
R;Atoda, H.; Ishikawa, M.; Yoshihara, E.; Sekiya, F.; Morita, T.
J. Biochem. 118, 965-973, 1995
A;Title: Blood coagulation factor IX-binding protein from the venom of Trimeresurus flav
A;Reference number: JC4329; MUID:96318509; PMID:8749314
A;Accession: JC4329
A;Molecule type: protein
A;Residues: 1-129 <ATO>
C;Comment: This protein binds calcium.
C;Superfamily: tetranectin; C-type lectin homology
C;Keywords: anticoagulant; blood coagulation; calcium binding; venom
F;2-127/Domain: C-type lectin homology <LCH>
F;2-13,30-127,102-119/Disulfide bonds: #status predicted

  Query Match             12.9%;  Score 143.5;  DB 2;  Length 129;
  Best Local Similarity   29.6%;  Pred. No. 1.4e-05;
  Matches   42;  Conservative   22;  Mismatches   47;  Indels   31;  Gaps    7;

Qy         20 CYKVIYFH---DTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFI-ENLLPSDGDF 75
              |||    :   | : |   |:||      || ||||||  |  : : : ||:  :
Db         13 CYKPFKLYKTWDDAERFCTEQAK------GGHLVSIESAGEADFVAQLVTENIQNTKSYV 66

Qy         76 WIGLR-RREEKQSNSTACQDLYAWTDGSISQFRNWYVDEPSCGSEVCVVMYHQPSAPAGI 134
              ||||| : :||| :|        |:|||   :  ||   |    |: |: :  :
Db         67 WIGLRVQGKEKQCSS-------EWSDGSSVSYENWIEAE----SKTCLGLEKETG----- 110

Qy        135 GGPYMFQWNDDRCNMKNNFICK 156
                    :| :  |  :| |:|:
Db        111 ----FRKWVNIYCGQQNPFVCE 128


RESULT 31
A32375
L-selectin precursor - mouse
N;Alternate names: lymph node homing receptor MEL-14; lymphocyte surface antigen Ly-22
C;Species: Mus musculus (house mouse)
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 21-Jan-2000
C;Accession: A32375; A35102; A40167; A60906
R;Lasky, L.A.; Singer, M.S.; Yednock, T.A.; Dowbenko, D.; Fennie, C.; Rodriguez, H.; Ngu
Cell 56, 1045-1055, 1989
A;Title: Cloning of a lymphocyte homing receptor reveals a lectin domain.
A;Reference number: A32375; MUID:89168433; PMID:2647302
A;Accession: A32375
A;Molecule type: mRNA
A;Residues: 1-372 <LAS>
A;Cross-references: GB:M25324; NID:g198803; PIDN:AAA39431.1; PID:g198804
R;Siegelman, M.H.; Cheng, I.C.; Weissman, I.L.; Wakeland, E.K.
Cell 61, 611-622, 1990
A;Title: The mouse lymph node homing receptor is identical with the lymphocyte cell surf
A;Reference number: A35102; MUID:90263086; PMID:1693096
A;Accession: A35102
A;Status: not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-372 <SIE1>
A;Cross-references: GB:M36005; NID:g199735; PIDN:AAA39722.1; PID:g199736; GB:M36058; NID
R;Siegelman, M.H.; van de Rijn, M.; Weissman, I.L.
Science 243, 1165-1172, 1989
A;Title: Mouse lymph node homing receptor cDNA clone encodes a glycoprotein revealing ta
A;Reference number: A40167; MUID:89162048; PMID:2646713
A;Accession: A40167
A;Molecule type: mRNA
A;Residues: 1-372 <SIE2>
A;Cross-references: GB:X14772; NID:g52942; PIDN:CAA32880.1; PID:g52943
A;Note: part of this sequence, including the amino end of the mature protein, was confirm
R;Siegelman, M.; Bond, M.W.; Gallatin, W.M.; St.John, T.; Smith, H.T.; Fried, V.A.; Weiss
Science 231, 823-829, 1986
A;Title: Cell surface molecule associated with lymphocyte homing is a ubiquitinated bran
A;Reference number: A60906; MUID:86122900; PMID:3003913
A;Accession: A60906
A;Molecule type: protein
A;Residues: 'X',40,'X',42,'XXX',46,'X',48,'XXXXXX',55-56,'X',58,'XXXXXX',65,'V',67,'XL',
C;Comment: This protein is ubiquitinated.
C;Function:
A;Description: binds with low affinity to oligosaccarides like heparan sulfate and sialy
ment of leucocytes to areas of inflammation, and with CD162 mediates neutrophil-neutroph
C;Superfamily: L-selectin; C-type lectin homology; complement factor H repeat homology;
C;Keywords: cell adhesion; duplication; glycoprotein; inflammation; phosphoprotein; surf
F;1-38/Domain: signal sequence #status predicted <SIG>
F;29-155/Domain: C-type lectin homology <LCH>
F;39-372/Product: L-selectin #status experimental <MAT>
F;39-331/Domain: extracellular #status predicted <EXT>
F;160-191/Domain: EGF homology <EGF>
F;197-254/Domain: complement factor H repeat homology <FH1>
F;259-316/Domain: complement factor H repeat homology <FH2>
F;332-355/Domain: transmembrane #status predicted <TMM>
F;356-372/Domain: intracellular #status predicted <INT>
F;60,104,216,246,278,308,320/Binding site: carbohydrate (Asn) (covalent) #status predict
F;364/Binding site: phosphate (Ser) (covalent) #status predicted

  Query Match             12.9%;  Score 143.5;  DB 1;  Length 372;
  Best Local Similarity   26.5%;  Pred. No. 4.8e-05;
  Matches   41;  Conservative   37;  Mismatches   50;  Indels   27;  Gaps    8;

Qy         26 FHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLP-SDGDFWIGLRRREE 84
              :| : : :|:|. |:: |:::   ||:|::: |   || ::|| || |   :|||:|:
Db         41 YHYSEKPMNWENARKFCKQNYTDLVAIQNKRE---IE-YLENTLPKSPYYYWIGIRK--- 93

Qy         85 KQSNSTACQDLYAW--TDGSIS-QFRNWYVDEPS--CGSEVCVVMYHQPSAPAGIGGPYM 139
                        :: |  |: ::: :  ||   ||:     | || :| :    :|
Db         94 -------IGKMWTWVGTNKTLTKEAENWGAGEPNNKKSKEDCVEIYIKRERDSG------ 140

Qy        140 FQWNDDRCNMKNNFICKYSDEKPAVPSREAEGEET 174
                :|||| |: :   :|  :  :|   :   |  ||
Db        141 -KWNDDACHKRKAALCYTASCQPGSCNGRGECVET 174


RESULT 32
T26655
hypothetical protein Y38E10A.e - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C;Accession: T26655
R;Wallis, J.
submitted to the EMBL Data Library, September 1999
A;Reference number: Z20252
A;Accession: T26655
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-463 <WIL>
A;Cross-references: EMBL:AL110484; NID:e1542205; PIDN:CAB54396.1; CESP:Y38E10A.e
A;Experimental source: clone Y38E10A
C;Genetics:
A;Gene: CESP:Y38E10A.e
A;Introns: 35/2; 107/3; 155/1; 198/1

  Query Match             12.9%;  Score 143.5;  DB 2;  Length 463;
  Best Local Similarity   23.8%;  Pred. No. 6.1e-05;
  Matches   45;  Conservative   26;  Mismatches   73;  Indels   45;  Gaps    8;

Qy         20 CYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLPSD---GDFW 76
              ||    ||:|:    : : :: | :: | | || | :| : |  |        |   |  |
Db        208 CYT---FHNTAS--TYTKGQKICEQECGNLASIHSANENRYIMTFGGRATKEDLLLGGMW 262
```

```
Qy          77 IGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEPSCGSEVCVVMYHQPSAPAGIGG 136
                                |:| | |||: :: |:  |  :    |||:| :  | |  :|
Db         263 --------------PADDVYNWVDGSLWEYENF--DPINVRDSVCVIMSNGDSRPIALG- 305

Qy         137 PYMFQWNDDRCNMKNNFICKYS-------------DEKPAVPS--REAEGEETELTTPV 180
                     |    |  : : :||                  |||  |    :    :|:|:|
Db         306 ----MWYSGECKNEYSVVCKRPAGIQCPPNPTIVPVTPMPAVQSFCNSSVMAPSEITSPH 361

Qy         181 LPEETQEED 189
                |    |
Db         362 FPYNYYNSD 370


RESULT 33
JC5058
bitiscetin alpha chain - puff adder
N;Alternate names: von Willebrand factor modulator protein
C;Species: Bitis arietans (puff adder)
C;Date: 31-Jan-1997 #sequence_revision 31-Jan-1997 #text_change 12-Feb-1999
C;Accession: JC5058; JC5916
R;Matsui, T.; Hamako, J.; Suzuki, M.; Hayashi, N.; Ito, M.; Makita, K.; Fujimura, Y.; Oz
submitted to JIPID, January 1997
A;Description: Complete amino acid sequence of bitiscetin, a novel von willebrand factor
A;Reference number: JC5058
A;Contents: snake venom
A;Accession: JC5058
A;Molecule type: protein
A;Residues: 1-131 <MAT>
A;Experimental source: snake venom
R;Matsui, T.; Hamako, J.; Suzuki, M.; Hayashi, N.; Ito, M.; Makita, K.; Fujimura, Y.; Oz
Res. Commun. Biochem. Cell Mol. Biol. 1, 271-284, 1997
A;Title: Complete amino acid sequence of bitiscetin, a novel von Willebrand factor modul
A;Reference number: JC5916
A;Accession: JC5916
A;Molecule type: protein
A;Residues: 1-131 <MA2>
A;Experimental source: venom
C;Comment: This protein is a modulator of a von Willebrand factor modulator.
C;Superfamily: tetranectin; C-type lectin homology
C;Keywords: venom
F;4-125/Domain: C-type lectin homology <LCH>

  Query Match             12.8%;  Score 142.5;  DB 2;  Length 131;
  Best Local Similarity   33.6%;  Pred. No. 1.7e-05;
  Matches   47;  Conservative   14;  Mismatches   48;  Indels   31;  Gaps    8;

Qy          20 CYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFI-ENLLPSDGDFWIG 78
               ||||      |     :|:|:: |  : | | ||:|::|   : |   : |     | |||
Db          15 CYKVFKKVGT-----WEDAEKFCVENSGHLASIDSKEEADFVTKLASQTLTKFVYDAWIG 69

Qy          79 LRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEPS-C-GSEVCVVMYHQPSAPAGIGG 136
               ||   : |  |        |||||   : |  ||||: | | :|     |
Db          70 LRDESKTQQCSP------QWTDGSSVVYEN--VDEPTKCFGLDV-----HTEYR------ 110

Qy         137 PYMFQWNDDRCNMKNNFICK 156
                    |  |   |  || ||||
Db         111 ----TWTDLPCGEKNPFICK 126


RESULT 34
S48719
phospholipase-A(2) receptor protein - mouse
C;Species: Mus musculus (house mouse)
C;Date: 07-May-1995 #sequence_revision 21-Jul-1995 #text_change 20-Jun-2000
C;Accession: S48719
R;Higashino, K.; Ishizaki, J.; Kishino, J.; Ohara, O.; Arita, H.
Eur. J. Biochem. 225, 375-382, 1994
A;Title: Structural comparison of phospholipase-A(2)-binding regions in phospholipase-A(
A;Reference number: S48719; MUID:95010128; PMID:7925459
A;Accession: S48719
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-1487 <HIG>
A;Cross-references: GB:D30779; NID:g1375042; PIDN:BAA06443.1; PID:g691754
C;Superfamily: phospholipase A2 receptor; C-type lectin homology; fibronectin type II re
F;181-222/Domain: fibronectin type II repeat homology <2F1>
F;380-503/Domain: C-type lectin homology <LCH>

  Query Match             12.7%;  Score 142;  DB 2;  Length 1487;
  Best Local Similarity   26.5%;  Pred. No. 0.00032;
  Matches   40;  Conservative   22;  Mismatches   51;  Indels   38;  Gaps    6;

Qy          13 RGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLPSD 72
               | |: | ||:     :    |:: :|  :|    || |:||  |||:  | | :  ::
Db         238 RNGSSRICYQ----FNLLSSLSWNQAHSSCLMQGGALLSIADEDEEDFIRKHLSKVVK-- 291

Qy          73 GDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVD-------EPSCGSEVCVVMY 125
                : |||| : :||          : |:||:   : ||  :       |  ||:   |
Db         292 -EVWIGLNQLDEKAG--------WQWSDGTPLSYLNWSQEITPGPFVEHHCGTLEVV--- 339

Qy         126 HQPSAPAGIGGPYMFQWNDDRCNMKNNFICK 156
                 ||          |    |     :|||
Db         340 ---SA----------AWRSRDCESTLPYICK 357


RESULT 35
LNMSER
IgE Fc receptor, low-affinity - mouse
N;Alternate names: Blast-2; CD23; Fc-epsilon-RII; lymphocyte IgE receptor
C;Species: Mus musculus (house mouse)
C;Date: 12-Feb-1993 #sequence_revision 28-Oct-1994 #text_change 22-Jun-1999
C;Accession: A43518; A33840
R;Gollnick, S.O.; Trounstine, M.L.; Yamashita, L.C.; Kehry, M.R.; Moore, K.W.
J. Immunol. 144, 1974-1982, 1990
A;Title: Isolation, characterization, and expression of cDNA clones encoding the mouse Fc
A;Reference number: A43518; MUID:90171598; PMID:2137845
A;Accession: A43518
A;Molecule type: mRNA
A;Residues: 1-331 <GOL>
A;Cross-references: GB:M34163; NID:g193242; PIDN:AAA37603.1; PID:g309227
R;Bettler, B.; Hofstetter, H.; Rao, M.; Yokoyama, W.M.; Kilchherr, F.; Conrad, D.H.
Proc. Natl. Acad. Sci. U.S.A. 86, 7566-7570, 1989
A;Title: Molecular structure and expression of the murine lymphocyte low-affinity recepto
A;Reference number: A33840; MUID:90017519; PMID:2529542
A;Accession: A33840
A;Molecule type: mRNA
A;Residues: 1-331 <BET>
A;Cross-references: GB:M99371; NID:g193245; PIDN:AAA74898.1; PID:g193246; GB:M27150
C;Comment: This receptor for the Fc portion of IgE is expressed in various hematopoietic
f B-cells.
C;Superfamily: IgE receptor II; C-type lectin homology
C;Keywords: B-cell; glycoprotein; immunoglobulin receptor; macrophage; tandem repeat; tra
F;1-25/Domain: intracellular #status predicted <INT>
F;14-22/Region: stop-transfer sequence
F;26-46/Domain: transmembrane #status predicted <TMM>
F;47-331/Domain: extracellular #status predicted <EXT>
F;66-86/Region: 21-residue repeat
F;87-107/Region: 21-residue repeat
F;108-128/Region: 21-residue repeat
F;129-149/Region: 21-residue repeat
F;186-305/Domain: C-type lectin homology <LCH>
F;65,114/Binding site: carbohydrate (Asn) (covalent) #status predicted

  Query Match             12.7%;  Score 141.5;  DB 1;  Length 331;
  Best Local Similarity   31.8%;  Pred. No. 6.3e-05;
  Matches   47;  Conservative   21;  Mismatches   37;  Indels   43;  Gaps    9;

Qy          17 QRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFW 76
               |: ||   ||   |::  : :|: ||    |:|||| |: ||  : : |     :  | |
Db         194 QQKCY---YFGKGSKQ--WIQARFACSDLQGRLVSIHSQKEQDFLMQHI-----NKKDSW 243

Qy          77 IGLRRREEKQSNSTACQDL-----YAWTDGSISQFRNWYVDEPSCG--SEVCVVMYHQPS 129
               |||              |||     : |:|||   : ||   ||: |   | ||:|
Db         244 IGL-------------QDLNMEGEFVWSDGSPVGYSNWNPGEPNNGGQGEDCVMM----- 285

Qy         130 APAGIGGPYMFQWNDDRC-NMKNNFICK 156
```

```
                  | |     ||||  | :  : ::|:
Db        286 --RGSG-----QWNDAFCRSYLDAWVCE 306
```

RESULT 36
LNRC3
lectin BRA3-2 precursor - barnacle (Megabalanus rosa)
C;Species: Megabalanus rosa
C;Date: 31-Dec-1988 #sequence_revision 09-Sep-1994 #text_change 16-Jul-1999
C;Accession: JC1504; A26094
R;Takamatsu, N.; Takeda, T.; Kojima, M.; Heishi, M.; Muramoto, K.; Kamiya, H.; Shiba, T.
Gene 128, 251-255, 1993
A;Title: Acorn barnacle Megabalanus rosa lectin (BRA-3): cDNA cloning, gene structure an
A;Reference number: JC1503; MUID:93292994; PMID:8514190
A;Accession: JC1504
A;Molecule type: mRNA
A;Residues: 1-162 <TAK>
R;Muramoto, K.; Kamiya, H.
Biochim. Biophys. Acta 874, 285-295, 1986
A;Title: The amino-acid sequence of a lectin of the acorn barnacle Megabalanus rosa.
A;Reference number: A26094
A;Accession: A26094
A;Molecule type: protein
A;Residues: 25-162 <MUR>
A;Note: 146-Arg was also found
C;Comment: This three galactose-binding lectin is isolated from the coelomic fluid.
C;Comment: This protein plays important roles in defense mechanisms and in development a
C;Comment: The molecule is a tetramer of identical chains.
C;Superfamily: tetranectin; C-type lectin homology
C;Keywords: hemolymph; homotetramer; lectin
F;1-24/Domain: signal sequence #status predicted <SIG>
F;25-162/Product: lectin BRA3-2 #status experimental <MAT>
F;26-150/Domain: C-type lectin homology <LCH>
F;26-39,56-150,125-142/Disulfide bonds: #status experimental
F;157/Disulfide bonds: interchain (to 160) #status experimental
F;160/Disulfide bonds: interchain (to 157) #status experimental

  Query Match 12.6%; Score 140.5; DB 1; Length 162;
  Best Local Similarity 24.4%; Pred. No. 3.3e-05;
  Matches 39; Conservative 26; Mismatches 60; Indels 35; Gaps 6;

```
Qy          5 LLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACR--RDGGQLVSIESEDEQKLIE 62
              : : :  | |       |    :  :: :: : :|: ||:    |  | :|:|: |   |
Db         19 ITTAECTCPGNLDWQEYDGHCYWASTYQVRWNDAQLACQTVHPGAYLATIQSQLENAFIS 78

Qy         63 KFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEPS------C 116
              : : |       ||||        |    :  | |::|  : |  |  : |:      |
Db         79 ETVSN-----NRLWIGL--------NDIDLEGHYVWSNGEATDFTYWSSNNPNNWENQDC 125

Qy        117 GSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICK 156
              |    || |   :           ||:|| ||   ||:||
Db        126 G----VVNYDTVTG----------QWDDDDCNKNKNFLCK 151
```

RESULT 37
S22124
L-selectin precursor - bovine
N;Alternate names: leucocyte cell adhesion molecule 1 (LECAM-1)
C;Species: Bos primigenius taurus (cattle)
C;Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 21-Jan-2000
C;Accession: S22124; A46531
R;Bosworth, B.T.
submitted to the EMBL Data Library, October 1991
A;Reference number: S22123
A;Accession: S22124
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-370 <BOS>
A;Cross-references: EMBL:X62882; NID:g515; PIDN:CAA44676.1; PID:g516
R;Walcheck, B.; White, M.; Kurk, S.; Kishimoto, T.K.; Jutila, M.A.
Eur. J. Immunol. 22, 469-476, 1992
A;Title: Characterization of the bovine peripheral lymph node homing receptor: a lectin
A;Reference number: A46531; MUID:92164727; PMID:1371468
A;Accession: A46531
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 39-79,'Q',81-151,'K' <WAL>
A;Note: sequence inconsistent with the nucleotide translation
A;Note: sequence extracted from NCBI backbone (NCBIN:85686, NCBIP:85687)
C;Superfamily: L-selectin; C-type lectin homology; complement factor H repeat homology;
F;29-155/Domain: C-type lectin homology <LCH>
F;160-191/Domain: EGF homology <EGF>
F;197-254/Domain: complement factor H repeat homology <FH1>
F;259-316/Domain: complement factor H repeat homology <FH2>

  Query Match 12.6%; Score 140; DB 2; Length 370;
  Best Local Similarity 25.3%; Pred. No. 9.7e-05;
  Matches 46; Conservative 35; Mismatches 57; Indels 44; Gaps 9;

```
Qy         26 FHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLP-SDGDFWIGLRRREE 84
              :| : | : :|:|:  || :   ||:|::: |   || ::   || |   :|||:|: |
Db         41 YHYSKRPMPWEKARAFCRENYTDLVAIQNKGE---IE-YLNKTLPFSRTYYWIGIRKVE- 95

Qy         85 KQSNSTACQDLYAW--TDGSIS-QFRNWYVDEPS--CGSEVCVVMYHQPSAPAGIGGPYM 139
                        :: |  |: |:: : :||   ||:     | || :| : :  :|
Db         96 ---------GVWTWVGTNKSLTEEAKNWGAGEPNNRKSKEDCVEIYIKRNKDSG------ 140

Qy        140 FQWNDDRCNMKNNFICKYSDEKPAVPSREAE----------------GEETELTTPVLP 182
              :|||| |:     :|  :  ||   |  :                 | | :  |  :|
Db        141 -KWNDDACHKAKTALCYTASCKPWSCSGHGQCVEVINNYTCNCDLGYYGPECQFVTQCVP 199

Qy        183 EE 184
               |
Db        200 LE 201
```

RESULT 38
LNHUPS
pulmonary surfactant protein A precursor (genomic clone) - human
N;Alternate names: alveolar proteinosis protein; pulmonary surfactant 32K apoprotein; pu
C;Species: Homo sapiens (man)
C;Date: 31-Dec-1990 #sequence_revision 31-Dec-1990 #text_change 16-Jul-1999
C;Accession: A24622; A43628
R;White, R.T.; Damm, D.; Miller, J.; Spratt, K.; Schilling, J.; Hawgood, S.; Benson, B.;
Nature 317, 361-363, 1985
A;Title: Isolation and characterization of the human pulmonary surfactant apoprotein gen
A;Reference number: A24622; MUID:86014366; PMID:2995821
A;Accession: A24622
A;Molecule type: DNA
A;Residues: 1-248 <WHI>
A;Cross-references: GB:M30838; NID:g190564; PIDN:AAA36510.1; PID:g190565
A;Note: the sequence in GenBank entry HUMPSAP, release 109.0, (PID:g190565) has the codo
A;Note: four nucleotide differences, producing amino acid differences at positions 45, 5
R;Haagsman, H.P.; White, R.T.; Schilling, J.; Lau, K.; Benson, B.J.; Golden, J.; Hawgood
Am. J. Physiol. 257, L421-L429, 1989
A;Title: Studies of the structure of lung surfactant protein SP-A.
A;Reference number: A43628; MUID:90119861; PMID:2610270
A;Accession: A43628
A;Molecule type: protein
A;Residues: 143-150;220-240;243-248 <HAA>
C;Comment: Pulmonary surfactant is a complex of phospholipids and proteins that lowers t
C;Comment: This protein is a sialoglycoprotein synthesized by alveolar type II cells. It
pendent on the presence of calcium ions.
C;Genetics:
A;Gene: GDB:SFTPA1; SFTP1; SP-A; SP-A1
A;Cross-references: GDB:119593; OMIM:178630
A;Map position: 10q22-10q23
A;Introns: 58/1; 98/1; 124/1
C;Superfamily: mannose-binding lectin; C-type lectin homology
C;Keywords: alveolar proteinosis; calcium; gaseous exchange; glycoprotein; hydroxylysine
F;1-20/Domain: signal sequence #status predicted <SIG>
F;21-248/Product: pulmonary surfactant protein A #status predicted <MAT>
F;28-100/Domain: collagenous #status predicted <COL>
F;127-246/Domain: C-type lectin homology <LCH>
F;26/Disulfide bonds: interchain #status experimental
F;30,33,36,42,57,63,76,79,82,91,97/Modified site: 4-hydroxyproline (Pro) #status predict

F;51,88/Modified site: 5-hydroxylysine (Lys) #status predicted
F;155-246,224-238/Disulfide bonds: #status experimental
F;207/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 12.5%; Score 139; DB 1; Length 248;
Best Local Similarity 25.2%; Pred. No. 7.4e-05;
Matches 36; Conservative 18; Mismatches 45; Indels 44; Gaps 4;

```
Qy         26 FHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLL----------PSDGDF 75
              |    : : |:  :||| | ||::    : :| : |  |::              || |||
Db        139 FSSNGQSITFDAIQEACARAGGRIAVPRNPEENEAIASFVKKYNTYAYVGLTEGPSPGDF 198

Qy         76 WIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEPS-CGSEVCVVMYHQPSAPAGI 134
                                  ::||:   : |||  ||:  | | || ||
Db        199 --------------------RYSDGTPVNYTNWYRGEPAGRGKEQCVEMYTDG------ 231

Qy        135 GGPYMFQWNDDRCNMKNNFICKY 157
                |||| |      ||::
Db        232 ------QWNDRNCLYSRLTICEF 248
```

RESULT 39
LNHUP6
pulmonary surfactant protein A precursor (clone 6A) - human
N;Alternate names: pulmonary surfactant 32K apoprotein; pulmonary surfactant-associated
C;Species: Homo sapiens (man)
C;Date: 31-Dec-1990 #sequence_revision 31-Dec-1990 #text_change 16-Jul-1999
C;Accession: A25720
R;Floros, J.; Steinbrink, R.; Jacobs, K.; Phelps, D.; Kriz, R.; Recny, M.; Sultzman, L.;
J. Biol. Chem. 261, 9029-9033, 1986
A;Title: Isolation and characterization of cDNA clones for the 35-kDa pulmonary surfacta
A;Reference number: A25720; MUID:86250832; PMID:3755136
A;Accession: A25720
A;Molecule type: mRNA
A;Residues: 1-248 <FLO>
A;Cross-references: GB:M13686; NID:g190669; PIDN:AAA60211.1; PID:g190670
A;Note: part of the sequence was confirmed by protein sequencing
A;Note: the amino end of the mature protein, which was not identified, is partially acet
A;Note: clones corresponding to two different proteins were sequenced. Cotranslational m
C;Genetics:
A;Gene: GDB:SFTPA1; SFTP1; SP-A; SP-A1
A;Cross-references: GDB:119593; OMIM:178630
A;Map position: 10q22-10q23
C;Superfamily: mannose-binding lectin; C-type lectin homology
C;Keywords: acetylated amino end; alveolar proteinosis; calcium; gaseous exchange; glycc
F;1-20/Domain: signal sequence #status predicted <SIG>
F;21-248/Product: pulmonary surfactant protein A #status predicted <MAT>
F;127-246/Domain: C-type lectin homology <LCH>
F;21/Modified site: acetylated amino end (Glu) (in mature form) #status predicted
F;30,33,36,42,54,57,63,76,79,82,91,97/Modified site: 4-hydroxyproline (Pro) #status pred
F;207/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 12.5%; Score 139; DB 1; Length 248;
Best Local Similarity 25.2%; Pred. No. 7.4e-05;
Matches 36; Conservative 18; Mismatches 45; Indels 44; Gaps 4;

```
Qy         26 FHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLL----------PSDGDF 75
              |    : : |:  :||| | ||::    : :| : |  |::              || |||
Db        139 FSSNGQSITFDAIQEACARAGGRIAVPRNPEENEAIASFVKKYNTYAYVGLTEGPSPGDF 198

Qy         76 WIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEPS-CGSEVCVVMYHQPSAPAGI 134
                                  ::||:   : |||  ||:  | | || ||
Db        199 --------------------RYSDGTPVNYTNWYRGEPAGRGKEQCVEMYTDG------ 231

Qy        135 GGPYMFQWNDDRCNMKNNFICKY 157
                |||| |      ||::
Db        232 ------QWNDRNCLYSRLTICEF 248
```

RESULT 40
LNHUP1
pulmonary surfactant protein A precursor (clone 1A) - human
N;Alternate names: pulmonary surfactant 32K apoprotein; pulmonary surfactant-associated
C;Species: Homo sapiens (man)
C;Date: 31-Dec-1990 #sequence_revision 31-Dec-1990 #text_change 16-Jul-1999
C;Accession: B25720
R;Floros, J.; Steinbrink, R.; Jacobs, K.; Phelps, D.; Kriz, R.; Recny, M.; Sultzman, L.;
J. Biol. Chem. 261, 9029-9033, 1986
A;Title: Isolation and characterization of cDNA clones for the 35-kDa pulmonary surfactar
A;Reference number: A25720; MUID:86250832; PMID:3755136
A;Accession: B25720
A;Molecule type: mRNA
A;Residues: 1-248 <FLO>
A;Cross-references: GB:K03475
A;Note: part of the sequence was confirmed by protein sequencing
A;Note: the amino end of the mature protein, which was not identified, is partially acety
A;Note: clones corresponding to two different proteins were sequenced. Cotranslational mc
C;Genetics:
A;Gene: GDB:SFTPA1; SFTP1; SP-A; SP-A1
A;Cross-references: GDB:119593; OMIM:178630
A;Map position: 10q22-10q23
C;Superfamily: mannose-binding lectin; C-type lectin homology
C;Keywords: acetylated amino end; alveolar proteinosis; calcium; gaseous exchange; glycoy
F;1-20/Domain: signal sequence #status predicted <SIG>
F;21-248/Product: pulmonary surfactant protein A #status predicted <MAT>
F;127-246/Domain: C-type lectin homology <LCH>
F;21/Modified site: acetylated amino end (Glu) (in mature form) #status predicted
F;30,33,36,42,54,57,63,76,79,82,91,97/Modified site: 4-hydroxyproline (Pro) #status predi
F;207/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 12.5%; Score 139; DB 1; Length 248;
Best Local Similarity 25.2%; Pred. No. 7.4e-05;
Matches 36; Conservative 18; Mismatches 45; Indels 44; Gaps 4;

```
Qy         26 FHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLL----------PSDGDF 75
              |    : : |:  :||| | ||::    : :| : |  |::              || |||
Db        139 FSSNGQSITFDAIQEACARAGGRIAVPRNPEENEAIASFVKKYNTYAYVGLTEGPSPGDF 198

Qy         76 WIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEPS-CGSEVCVVMYHQPSAPAGI 134
                                  ::||:   : |||  ||:  | | || ||
Db        199 --------------------RYSDGTPVNYTNWYRGEPAGRGKEQCVEMYTDG------ 231

Qy        135 GGPYMFQWNDDRCNMKNNFICKY 157
                |||| |      ||::
Db        232 ------QWNDRNCLYSRLTICEF 248
```

RESULT 41
LNFHLS
lectin precursor - flesh fly (Sarcophaga peregrina)
C;Species: Sarcophaga peregrina
C;Date: 31-Dec-1988 #sequence_revision 31-Dec-1988 #text_change 16-Jun-2000
C;Accession: S07759; S41119; A25736; B25736
R;Kobayashi, A.; Hirai, H.; Kubo, T.; Ueno, K.; Nakanishi, Y.; Natori, S.
Biochim. Biophys. Acta 1009, 244-250, 1989
A;Title: Cloning and in vitro transcription of the Sarcophaga lectin gene.
A;Reference number: S07759; MUID:90089397; PMID:2480809
A;Accession: S07759
A;Molecule type: DNA
A;Residues: 1-38 <KOB>
A;Cross-references: EMBL:X16659; NID:g10271; PIDN:CAA34645.1; PID:g10272
R;Matsui, M.; Kobayashi, A.; Kubo, T.; Natori, S.
Eur. J. Biochem. 219, 449-454, 1994
A;Title: Purification and characterization of ATBP, a novel protein that binds to A/T st
A;Reference number: S41119; MUID:94139722; PMID:8307011
A;Accession: S41119
A;Status: preliminary; translation not shown
A;Molecule type: DNA
A;Residues: 1-283 <MAT>
A;Cross-references: EMBL:D14870; NID:g287440; PIDN:BAA03586.1; PID:g391905
R;Takahashi, H.; Komano, H.; Kawaguchi, N.; Kitamura, N.; Nakanishi, S.; Natori, S.
J. Biol. Chem. 260, 12228-12233, 1985
A;Title: Cloning and sequencing of cDNA of Sarcophaga peregrina humoral lectin induced or
A;Reference number: A25736; MUID:86008294; PMID:2413021
A;Accession: A25736
A;Molecule type: mRNA

```
A;Residues: 1-283 <TAK>
A;Cross-references: GB:M11673; NID:g161264; PIDN:AAA29983.1; PID:g161265
A;Accession: B25736
A;Molecule type: protein
A;Residues: 24-32;164-171;209-214 <TAK2>
A;Note: the initiator Met may be either residue 1 or residue 5
C;Comment: This lectin is induced in third instar larvae when the body wall is injured a
C;Genetics:
A;Introns: 43/3
C;Superfamily: Sarcophaga lectin; C-type lectin homology
C;Keywords: hemolymph; lectin
F;1-23/Domain: signal sequence #status predicted <SIG>
F;24-283/Product: lectin #status experimental <LTN>
F;24-157/Domain: C-type lectin homology <LCH>

  Query Match             12.5%;  Score 139;  DB 1;  Length 283;
  Best Local Similarity   23.4%;  Pred. No. 8.7e-05;
  Matches   41;  Conservative   39;  Mismatches   79;  Indels   16;  Gaps    6;

Qy          25 YFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFWIGLRRREE 84
               |  :|  : |: :|   | |    |||:||| |:   |   :: ::    : |:|     :
Db          36 YLIETELKYNWHQAWHECARHDQQLVTIESADKNNAIIDLVKRVVGKSHNLWLG---GND 92

Qy          85 KQSNSTACQDLYAWT-DGSISQFRNWYVDEPS--CGSEVCVVMYHQPSAPAGIGGPYMFQ 141
               : |:|      : |:  |     |  | : |       | || ::   : |        ::|
Db          93 EYSSSRDYGRPFFWSPTGQAFSFAYWSENNPDNYKHQEHCVHIW--DTKP-------LYQ 143

Qy         142 WNDDRCNMKNNFICKYSDEKPAV-PSREAEGEETELTTPVLPEETQEEDAKKTFK 195
               |||: ||:|  :||: :  :     : : :  |  ::|   :  |  :  ||:: :
Db         144 WNDNDCNVKMGYICEPNHFRETYDQALKQKCEAIKITNSKISTEFDQLHAKQSLE 198


RESULT 42
JC4690
coagulation factor IX/factor X-binding protein chain A precursor - habu
C;Species: Trimeresurus flavoviridis (habu)
C;Date: 16-Aug-1996 #sequence_revision 16-Aug-1996 #text_change 20-Jun-2000
C;Accession: JC4690; A39332
R;Matsuzaki, R.; Yoshiara, E.; Yamada, M.; Shima, K.; Atoda, H.; Morita, T.
Biochem. Biophys. Res. Commun. 220, 382-387, 1996
A;Title: cDNA cloning of IX/X-BP, a heterogeneous two-chain anticoagulant protein from s
A;Reference number: JC4690; MUID:96184662; PMID:8645314
A;Accession: JC4690
A;Molecule type: mRNA
A;Residues: 1-152 <MAT1>
A;Cross-references: DDBJ:D83331; NID:g1402639; PIDN:BAA11887.1; PID:g1402640
A;Experimental source: venom
R;Atoda, H.; Hyuga, M.; Morita, T.
J. Biol. Chem. 266, 14903-14911, 1991
A;Title: The primary structure of coagulation factor IX/factor X-binding protein isolate
otein, tetranectin, and lymphocyte Fc epsilon receptor for immunoglobulin E.
A;Reference number: A39332; MUID:91332000; PMID:1831197
A;Accession: A39332
A;Status: preliminary
A;Molecule type: protein
A;Residues: 24-152 <ATO>
C;Superfamily: tetranectin; C-type lectin homology
C;Keywords: anticoagulant; blood coagulation; glycoprotein; hemolymph; lectin
F;1-23/Domain: signal sequence #status predicted <SIG>
F;24-152/Product: factor IX/X binding protein chain A #status predicted <MAT>
F;25-150/Domain: C-type lectin homology <LCH>
F;25-36,53-150,125-142/Disulfide bonds: #status predicted

  Query Match             12.4%;  Score 138.5;  DB 2;  Length 152;
  Best Local Similarity   28.4%;  Pred. No. 4.7e-05;
  Matches   40;  Conservative   23;  Mismatches   49;  Indels   29;  Gaps    7;

Qy          20 CYKVIYFHDTSRRLNFEEAKEAC--RRDGGQLVSIESEDEQKLIEKFI-ENLLPSDGDFW 76
               |||    : |     :|:|:  |  :  | ||||||  |   : : : :|:   |   |
Db          36 CYKAFEKYKT-----WEDAERVCTEQAKGAHLVSIESSGEADFVAQLVTQNMKRLDFYIW 90

Qy          77 IGLR-RREEKQSNSTACQDLYAWTDGSISQFRNWYVDEPSCGSEVCVVMYHQPSAPAGIG 135
               |||| : : || ||        |:|||   : ||   |    |: |: :  :
Db          91 IGLRVQGKVKQCNS-------EWSDGSSVSYENWIEAE----SKTCLGLEKETD------ 133

Qy         136 GPYMFQWNDDRCNMKNNFICK 156
                  :| :  |  :| |:|:
Db         134 ---FRKWVNIYCGQQNPFVCE 151


RESULT 43
JC4031
tetranectin precursor - mouse
C;Species: Mus musculus (house mouse)
C;Date: 13-Jun-1995 #sequence_revision 14-Jul-1995 #text_change 20-Aug-1999
C;Accession: JC4031
R;Sorensen, C.B.; Berglund, L.; Petersen, T.E.
Gene 152, 243-245, 1995
A;Title: Cloning of a cDNA encoding murine tetranectin.
A;Reference number: JC4031; MUID:95137396; PMID:7835708
A;Accession: JC4031
A;Molecule type: mRNA
A;Residues: 1-202 <SOR>
A;Cross-references: EMBL:X79199; NID:g671561; PIDN:CAA55791.1; PID:g671562
A;Experimental source: lung
C;Comment: This protein binds plasminogen, and may play a role in invasive cancer.
C;Superfamily: tetranectin; C-type lectin homology
F;1-21/Domain: signal sequence #status predicted <SIG>
F;22-202/Product: tetranectin #status predicted <MAT>
F;71-197/Domain: C-type lectin homology <LCH>

  Query Match             12.4%;  Score 138.5;  DB 2;  Length 202;
  Best Local Similarity   26.3%;  Pred. No. 6.5e-05;
  Matches   41;  Conservative   30;  Mismatches   56;  Indels   29;  Gaps    7;

Qy           9 QPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENL 68
               | ||  ||:     :: |   ::   | || | |   || | : :|| | : : ::  :
Db          68 QTVCLKGTKVNLKCLLAF---TQPKTFHEASEDCISQGGTLGTPQSELENEALFEYARHS 124

Qy          69 LPSDGDFWIGLRRREEKQSNSTACQDLYAWTD--GSISQFRNWYVD---EPSCG-SEVCV 122
               : :| : |:||        |  | :  || |  | :  ::||  :    :|  | :| | |
Db         125 VGNDANIWLGL--------NDMAAEG--AWVDMTGGLLAYKNWETEITTQPDGGKAENC- 173

Qy         123 VMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYS 158
                     | : |    :| | ||  :  :||:::
Db         174 ---------AALSGAANGKWFDKRCRDQLPYICQFA 200


RESULT 44
T29200
hypothetical protein T03F1.10 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C;Accession: T29200
R;Du, Z.; Le, T.T.
submitted to the EMBL Data Library, February 1997
A;Description: The sequence of C. elegans cosmid T03F1.
A;Reference number: Z20586
A;Accession: T29200
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-280 <DUZ>
A;Cross-references: EMBL:U88169; PIDN:AAB42238.1; GSPDB:GN00019; CESP:T03F1.10
A;Experimental source: strain Bristol N2; clone T03F1
C;Genetics:
A;Gene: CESP:T03F1.10
A;Map position: 1
A;Introns: 4/2; 134/1; 153/3; 189/1; 205/3; 245/3

  Query Match             12.4%;  Score 138;  DB 2;  Length 280;
  Best Local Similarity   28.6%;  Pred. No. 0.0001;
  Matches   42;  Conservative   18;  Mismatches   55;  Indels   32;  Gaps    6;

Qy          20 CYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLPS-------- 71
               |||      | :  || :|: ||: ||| | || |  |       |:  |  :
Db         151 CYKT-----TVQFTNFNDARSACQADGGDLASIHSIAE----NAFLVGLSKAGITNKDKD 201
```

```
Qy          72 -DGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEP-SCGSEVCVVMYHQPS 129
               : | ||||  :  |          : |||||:  : ||   || :   |    :   |
Db         202 WNDDVWIGLVYQNSK----------WQWTDGSVVNYVNWGDGEPNNMNKEWWTALVADPH 251

Qy         130 APAGIGGPYMFQWNDDRCNMKNNFICK 156
                     |    :||:   : :  |:||
Db         252 EGKNSEGT---RWNNVPQDDQRAFLCK 275


RESULT 45
S34198
IgE Fc receptor II, low-affinity - rat
N;Alternate names: CD23; lymphocyte IgE receptor
C;Species: Rattus norvegicus (Norway rat)
C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 22-Jun-1999
C;Accession: S34198
R;Flores-Romo, L.; Shield, J.; Humbert, Y.; Graber, P.; Aubry, J.P.; Gauchat, J.F.; Ayal
submitted to the EMBL Data Library, June 1993
A;Description: Inhibition of an in vivo antigen-specific IgE response by antibodies to C
A;Reference number: S34198
A;Accession: S34198
A;Molecule type: mRNA
A;Residues: 1-309 <FLO>
A;Cross-references: EMBL:X73579; NID:g313672; PIDN:CAA51981.1; PID:g313673
C;Superfamily: IgE receptor II; C-type lectin homology
C;Keywords: B-cell; glycoprotein; immunoglobulin receptor; macrophage; tandem repeat; tr
F;1-25/Domain: intracellular #status predicted <INT>
F;14-22/Region: stop-transfer sequence
F;24-46/Domain: transmembrane #status predicted <TMM>
F;47-309/Domain: extracellular #status predicted <EXT>
F;126-309/Product: soluble IgE-binding factor (29K) #status predicted <IGI>
F;149-309/Product: soluble IgE-binding factor (25-27K) #status predicted <BF1>
F;164-283/Domain: C-type lectin homology <LCH>
F;192-283,260-274/Disulfide bonds: #status predicted

  Query Match             12.3%;  Score 137.5;  DB 1;  Length 309;
  Best Local Similarity   31.1%;  Pred. No. 0.00013;
  Matches   46;  Conservative   21;  Mismatches   38;  Indels   43;  Gaps    9;

Qy          17 QRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFW 76
               |: ||   || : |::  : :||  |    |:|||| |: ||  : : |     :  : |
Db         172 QQKCY---YFGEGSKQ--WIQAKFTCSDLEGRLVSIHSQKEQDFLMQHI-----NKKESW 221

Qy          77 IGLRRREEKQSNSTACQDL-----YAWTDGSISQFRNWYVDEPSCG--SEVCVVMYHQPS 129
               |||           |||     : | |||   : ||   ||: |   | ||:|
Db         222 IGL------------QDLNMEGEFVWPDGSPVGYSNWNPGEPNNGGQGEDCVMM----- 263

Qy         130 APAGIGGPYMFQWNDDRC-NMKNNFICK 156
                  | |     ||||  | :  : ::|:
Db         264 --RGSG-----QWNDAFCRSYLDAWVCE 284


RESULT 46
JC5377
L-selectin precursor - hamadryas baboon
C;Species: Papio hamadryas (hamadryas baboon)
C;Date: 02-Jun-1997 #sequence_revision 12-Sep-1997 #text_change 21-Jan-2000
C;Accession: JC5377; PC4315
R;Tsurushita, N.; Fu, H.; Berg, E.L.
Gene 181, 219-220, 1996
A;Title: PCR cloning of the cDNA encoding baboon L-selectin.
A;Reference number: JC5377; MUID:97128794; PMID:8973334
A;Accession: JC5377
A;Molecule type: mRNA
A;Residues: 1-372 <TSU1>
A;Cross-references: GB:U52074; NID:g1326148; PIDN:AAB40903.1; PID:g1326149
A;Accession: PC4315
A;Molecule type: protein
A;Residues: 37-43;142-148 <TSU2>
C;Comment: This receptor is involved in the initial adhesive interaction between lympho
 sites of inflammation.
C;Superfamily: L-selectin; C-type lectin homology; complement factor H repeat homology;
F;1-38/Domain: signal sequence #status predicted <SIG>
F;29-155/Domain: C-type lectin homology <LCH>
F;39-372/Product: L-selectin #status predicted <MAT>
F;39-157/Domain: calcium-binding #status predicted <CAB>
F;160-191/Domain: EGF homology <EGF>
F;197-254/Domain: complement factor H repeat homology <FH1>
F;259-316/Domain: complement factor H repeat homology <FH2>
F;333-355/Domain: transmembrane #status predicted <TMM>
F;356-372/Domain: intracellular #status predicted <INT>

  Query Match             12.3%;  Score 137.5;  DB 2;  Length 372;
  Best Local Similarity   28.0%;  Pred. No. 0.00016;
  Matches   44;  Conservative   31;  Mismatches   53;  Indels   29;  Gaps    9;

Qy          26 FHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLP-SDGDFWIGLRRREE 84
               :| :   :|:::|:  || :   ||:|::: |   || ::|  || |   :|||:|:
Db          41 YHYSENPMNWQKARRFCRENYTDLVAIQNKAE---IE-YLEKTLPFSPSYYWIGIRK--- 93

Qy          85 KQSNSTACQDLYAW--TDGSISQ-FRNWYVDEPS--CGSEVCVVMYHQPSAPAGIGGPYM 139
                        :: |  |: |::|   ||   ||:     | || :| :    ||
Db          94 -------IGGIWTWVGTNKSLTQEAENWGDGEPNNKKNKEDCVEIYIKRKKDAG------ 140

Qy         140 FQWNDDRCNMKNNFICKYSDEKPAVPSREAEGEETEL 176
                :|||| |:     :|  :  :|   |    ||  |:
Db         141 -KWNDDACHKPKAALCYTASCQPW--SCSGHGECVEI 174


RESULT 47
JC4892
L-selectin precursor - rabbit
C;Species: Oryctolagus cuniculus (domestic rabbit)
C;Date: 04-Sep-1998 #sequence_revision 04-Sep-1998 #text_change 21-Jan-2000
C;Accession: JC4892
R;Qian, J.; Huang, X.; Marks, R.M.
Biochem. Biophys. Res. Commun. 225, 406-412, 1996
A;Title: Cloning of the cDNA for rabbit L-selectin and expression of recombinant protein
A;Reference number: JC4892; MUID:96354800; PMID:8753776
A;Accession: JC4892
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-376 <QIA>
A;Cross-references: GB:U26535; NID:g847787; PIDN:AAA67896.1; PID:g847788
C;Comment: This protein involved in leukocyte-endothelial adhesion; it mediates adhesion
C;Superfamily: L-selectin; C-type lectin homology; complement factor H repeat homology; E
F;1-37/Domain: signal sequence #status predicted <SIG>
F;29-155/Domain: C-type lectin homology <LCH>
F;38-376/Product: L-selectin #status predicted <MAT>
F;160-191/Domain: EGF homology <EGF>
F;197-254/Domain: complement factor H repeat homology <FH1>
F;259-316/Domain: complement factor H repeat homology <FHR>

  Query Match             12.3%;  Score 137.5;  DB 2;  Length 376;
  Best Local Similarity   27.4%;  Pred. No. 0.00016;
  Matches   43;  Conservative   33;  Mismatches   52;  Indels   29;  Gaps    9;

Qy          26 FHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLP-SDGDFWIGLRRREE 84
               :| : : :|:| |:: || :   ||:|::: |   || ::|  || |   :|||:|:
Db          41 YHYSEKPMNWERARKFCRENYTDLVAIQNKGE---IE-YLEKTLPFSRSYYWIGIRK--- 93

Qy          85 KQSNSTACQDLYAW--TDGSI-SQFRNWYVDEPS--CGSEVCVVMYHQPSAPAGIGGPYM 139
                        ::: |  |: |: ::  ||   ||:     | || :| :    :|
Db          94 -------IGNIWTWVGTNKSLTAEAENWGEGEPNNKKTKEDCVEIYIKRLRDSG------ 140

Qy         140 FQWNDDRCNMKNNFICKYSDEKPAVPSREAEGEETEL 176
                :|||| |  :   :|  :   |   |    ||  |:
Db         141 -KWNDDSCQKRKAALCYTASCHPG--SCSGHGECVEV 174


RESULT 48
I51921
pulmonary surfactant-associated protein A1 - human
N;Alternate names: SP-A1
C;Species: Homo sapiens (man)
```

```
C;Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 20-Aug-1999
C;Accession: I51921
R;Katyal, S.L.; Singh, G.; Locker, J.
Am. J. Respir. Cell Mol. Biol. 6, 446-452, 1992
A;Title: Characterization of a second human pulmonary surfactant-associated protein SP-A
A;Reference number: I51921; MUID:92198680; PMID:1372511
A;Accession: I51921
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-248 <RES>
A;Cross-references: GB:M68519; NID:g338048; PIDN:AAA60319.1; PID:g338049
C;Genetics:
A;Gene: GDB:SFTPA1; SFTP1; SP-A; SP-A1
A;Cross-references: GDB:119593; OMIM:178630
A;Map position: 10q22-10q23
A;Introns: 58/1; 98/1; 124/1
C;Superfamily: mannose-binding lectin; C-type lectin homology
F;127-246/Domain: C-type lectin homology <LCH>

  Query Match             12.3%;  Score 137;  DB 2;  Length 248;
  Best Local Similarity   25.2%;  Pred. No. 0.00011;
  Matches   36;  Conservative   17;  Mismatches   46;  Indels   44;  Gaps    4;

Qy         26 FHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLL----------PSDGDF 75
              |    : : |:  :||| | ||::    : :| : |  |::              || |||
Db        139 FSSNGQSITFDAIQEACARAGGRIAVPRNPEENEAIASFVKKYNTYAYVGLTEGPSPGDF 198

Qy         76 WIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEPS-CGSEVCVVMYHQPSAPAGI 134
                                  ::||:   : |||  ||:  | | || ||
Db        199 --------------------RYSDGTPVNYTNWYRGEPAGRGKEQCVEMYTDG------ 231

Qy        135 GGPYMFQWNDDRCNMKNNFICKY 157
                |||| |      || :
Db        232 ------QWNDRNCLYSRLTICDF 248


RESULT 49
T29754
hypothetical protein B0218.6 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C;Accession: T29754
R;Woessner, J.; Wohldmann, P.
submitted to the EMBL Data Library, May 1996
A;Description: The sequence of C. elegans cosmid B0218.
A;Reference number: Z20678
A;Accession: T29754
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-308 <WOE>
A;Cross-references: EMBL:U58752; PIDN:AAB00668.1; GSPDB:GN00022; CESP:B0218.6
A;Experimental source: strain Bristol N2; clone B0218
C;Genetics:
A;Gene: CESP:B0218.6
A;Map position: 4
A;Introns: 32/2; 162/1; 181/3; 217/1; 233/3; 273/3

  Query Match             12.2%;  Score 136.5;  DB 2;  Length 308;
  Best Local Similarity   26.9%;  Pred. No. 0.00016;
  Matches   47;  Conservative   19;  Mismatches   66;  Indels   43;  Gaps    8;

Qy          5 LLSGQPVCRGGTQRPCYKVIYFHDTSRRLN-------FEEAKEACRRDGGQLVSIESEDE 57
              |:| |  |  |         ||    || ||       |: |:| |    | |||| :  :
Db         15 LVSAQ--CGPG-------AIYQASGSRCLNVFRTASSFQSAEEICATLSGHLVSIHNTID 65

Qy         58 QKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDL-----YAWTDGSISQFRNWYVD 112
                 |    :     ||  |:| :         |   |:     : |||||   ::|: |
Db         66 NTFISSQAQKYF--DGGVWLGAK---------TTAPDVTNPLNWYWTDGSNFDYQNYKVG 114

Qy        113 EPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDE-KPAVPS 166
              :||   :   :     :|          :|    |  :  ||| ||     |  ||
Db        115 QPSPQGDTACMQLQTGTA----------KWQTTNCTAQLPFICSYSSSVTPTCPS 159


RESULT 50
A37194
pancreatic thread protein precursor - bovine
C;Species: Bos primigenius taurus (cattle)
C;Date: 31-Jan-1992 #sequence_revision 31-Jan-1992 #text_change 20-Aug-1999
C;Accession: A37194; A53897
R;de la Monte, S.M.; Ozturk, M.; Wands, J.R.
J. Clin. Invest. 86, 1004-1013, 1990
A;Title: Enhanced expression of an exocrine pancreatic protein in Alzheimer's disease an
A;Reference number: A37194; MUID:90368981; PMID:2394826
A;Accession: A37194
A;Molecule type: mRNA
A;Residues: 1-175 <DEL>
A;Cross-references: GB:M59794; NID:g163648; PIDN:AAA30750.1; PID:g163649
R;Cai, L.; Harris, W.R.; Marshak, D.R.; Gross, J.; Crabb, J.W.
J. Protein Chem. 9, 623-632, 1990
A;Title: Structural analysis of bovine pancreatic thread protein.
A;Reference number: A53897; MUID:91197388; PMID:2085387
A;Accession: A53897
A;Molecule type: protein
A;Residues: 38-138;141-175 <CAI>
C;Comment: The purified protein undergoes a reversible globule-fibril transformation and
C;Superfamily: tetranectin; C-type lectin homology
C;Keywords: disulfide bond; extracellular protein; pancreas
F;38-138/Product: pancreatic thread protein chain A #status experimental <ACH>
F;40-171/Domain: C-type lectin homology <LCH>
F;141-175/Product: pancreatic thread protein chain B #status experimental <BCH>

  Query Match             12.2%;  Score 136;  DB 2;  Length 175;
  Best Local Similarity   24.1%;  Pred. No. 9.1e-05;
  Matches   41;  Conservative   23;  Mismatches   64;  Indels   42;  Gaps    7;

Qy          5 LLSGQPVCRGGT---QRPCYKVIYFHDTSRRLNFEEAKEAC-RRDGGQLVSIESEDEQKL 60
              | | :  |  |:   :  || :     |     : :|  || :|  | |||: |  |:
Db         33 LPSARISCPSGSMAYRSHCYALFKTPKT-----WMDADIACQKRPSGHLVSVLSGAEESF 87

Qy         61 IEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEPSCGSEV 120
               :   : | | :  | ||||    :    |· |    : |    :  :  |  |
Db         88 VASLVRNNLNTQSDIWIGL---HDPTEGSEANAGGWEWISNDVLNYVAWETD-------- 136

Qy        121 CVVMYHQPSAPAGIGGP----------YMFQWNDDRCNMKNNFICKYSD 159
                    || |  |            |: :| |  ||:   ::||::|
Db        137 ----------PAAISSPGYCGSLSRSSGYL-KWRDHNCNLNLPYVCKFTD 175


Search completed: September  9, 2004, 22:39:38
Job time : 42 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: September 9, 2004, 22:23:47 ; Search time 23 Seconds
(without alignments)
466.367 Million cell updates/sec

Title: US-09-887-855-2_COPY_22_227
Perfect score: 1115
Sequence: 1 ATGRLLSGQPVCRGGTQRPC.....EEDAKKTFKESREAALNLAY 206

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 50 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	574	51.5	273	1	CHOD_MOUSE	Q9cxm0 mus musculu
2	561.5	50.4	273	1	CHOD_HUMAN	Q9h9p2 homo sapien
3	182	16.3	1456	1	MANR_HUMAN	P22897 homo sapien
4	177	15.9	1268	1	PGCN_MOUSE	P55066 mus musculu
5	176	15.8	1321	1	PGCN_HUMAN	O14594 homo sapien
6	174.5	15.7	3381	1	PGCV_BOVIN	P81282 bos taurus
7	174	15.6	1257	1	PGCN_RAT	P55067 rattus norv
8	174	15.6	2738	1	PGCV_RAT	Q9erb4 rattus norv
9	174	15.6	3358	1	PGCV_MOUSE	Q62059 mus musculu
10	174	15.6	3396	1	PGCV_HUMAN	P13611 homo sapien
11	171	15.3	3562	1	PGCV_CHICK	Q90953 gallus gall
12	165	14.8	643	1	CD93_RAT	Q9et61 rattus norv
13	158.5	14.2	2364	1	PGCA_BOVIN	P13608 bos taurus
14	158.5	14.2	2415	1	PGCA_HUMAN	P16112 homo sapien
15	155.5	13.9	2333	1	PGCA_CANFA	Q28343 canis famil
16	154.5	13.9	612	1	LEM2_MOUSE	Q00690 mus musculu
17	153.5	13.8	644	1	CD93_MOUSE	O89103 mus musculu
18	153.5	13.8	2124	1	PGCA_RAT	P07897 rattus norv
19	152	13.6	912	1	PGCB_BOVIN	Q28062 bos taurus
20	151.5	13.6	652	1	CD93_HUMAN	Q9npy3 homo sapien
21	151	13.5	197	1	CLE1_HUMAN	O75596 homo sapien
22	149	13.4	321	1	FCE2_HUMAN	P06734 homo sapien
23	148.5	13.3	2132	1	PGCA_MOUSE	Q61282 mus musculu
24	147	13.2	883	1	PGCB_MOUSE	Q61361 mus musculu
25	146	13.1	2109	1	PGCA_CHICK	P07898 gallus gall
26	145	13.0	158	1	LECG_TRIST	Q9ygp1 trimeresuru
27	145	13.0	883	1	PGCB_RAT	P55068 rattus norv
28	144.5	13.0	173	1	LEC2_MEGRO	P17346 megabalanus
29	144.5	13.0	372	1	LEM1_RAT	P30836 rattus norv
30	143.5	12.9	372	1	LEM1_MOUSE	P18337 mus musculu
31	141.5	12.7	331	1	FCE2_MOUSE	P20693 mus musculu
32	141.5	12.7	549	1	LEM2_RAT	P98105 rattus norv
33	140.5	12.6	162	1	LEC3_MEGRO	P07439 megabalanus
34	140	12.6	370	1	LEM1_BOVIN	P98131 bos taurus
35	139	12.5	248	1	PSPA_HUMAN	P07714 homo sapien
36	139	12.5	283	1	LECA_SARPE	P05047 sarcophaga
37	138.5	12.4	152	1	IXA_TRIFL	P23806 trimeresuru
38	138.5	12.4	202	1	TETN_MOUSE	P43025 mus musculu
39	137.5	12.3	372	1	LEM1_MACMU	Q95198 macaca mula
40	137.5	12.3	372	1	LEM1_PAPHA	Q28768 papio hamad
41	136	12.2	175	1	LITH_BOVIN	P23132 bos taurus
42	135	12.1	165	1	LIT1_MOUSE	P43137 mus musculu
43	134.5	12.1	132	1	ACAL_ANSAN	P83300 anser anser
44	134.5	12.1	372	1	LEM1_PONPY	Q95235 pongo pygma
45	134.5	12.1	485	1	LEM2_BOVIN	P98107 bos taurus
46	134	12.0	166	1	TETN_CARSP	P26258 carcharhinu
47	133.5	12.0	133	1	ECHA_ECHCA	P81017 echis carin
48	133.5	12.0	311	1	LECI_HUMAN	P07307 homo sapien
49	132.5	11.9	372	1	LEM1_HUMAN	P14151 homo sapien
50	132.5	11.9	372	1	LEM1_PANTR	Q95237 pan troglod

ALIGNMENTS

RESULT 1
CHOD_MOUSE
ID CHOD_MOUSE STANDARD; PRT; 273 AA.
AC Q9CXM0; Q8VI31;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Chondrolectin precursor (Transmembrane protein MT75).
GN CHODL.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J;
RA Weng L., Smits P., Hubner R., Wouters J., Merregaert J.;
RT "Mt75, a low expressed c-type lectin gene involving in
RT chondrogenesis.";
RL Submitted (OCT-2000) to the EMBL/GenBank/DDBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Embryonic head;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
CC -!- SUBCELLULAR LOCATION: Type I membrane protein (Potential).
CC -!- SIMILARITY: Contains 1 C-type lectin family domain.
CC ---

```
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; AF311699; AAL50354.1; -.
DR   EMBL; AK014255; BAB29226.1; -.
DR   HSSP; P22897; 1EGG.
DR   MGD; MGI:2179069; Chodl.
DR   InterPro; IPR001304; Lectin_C.
DR   Pfam; PF00059; lectin_c; 1.
DR   SMART; SM00034; CLECT; 1.
DR   PROSITE; PS00615; C_TYPE_LECTIN_1; FALSE_NEG.
DR   PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
KW   Lectin; Transmembrane; Glycoprotein; Signal.
FT   SIGNAL        1     21       POTENTIAL.
FT   CHAIN        22    273       CHONDROLECTIN.
FT   DOMAIN       22    216       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    217    237       POTENTIAL.
FT   DOMAIN      238    273       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN       35    179       C-TYPE LECTIN.
FT   CARBOHYD     86     86       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CONFLICT     24     24       V -> W (IN REF. 2).
FT   CONFLICT    179    179       T -> K (IN REF. 2).
SQ   SEQUENCE   273 AA;  30303 MW;  E052D933F244F4C7 CRC64;

  Query Match             51.5%;  Score 574;  DB 1;  Length 273;
  Best Local Similarity   57.4%;  Pred. No. 2.5e-44;
  Matches  109;  Conservative   26;  Mismatches   41;  Indels   14;  Gaps    5;

Qy          4 RLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEK 63
              |::||| ||    : ||||: |||: | |::|:||: ||  :|| |:|:|:| ||||||
Db         23 RVVSGQKVCFADVKHPCYKMAYFHELSSRVSFQEARLACESEGGVLLSLENEAEQKLIES 82

Qy         64 FIENLLP-----SDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEPSCGS 118
               ::||       |||||||||  :  |: | || ||| |:||| ||||||| |||||||
Db         83 MLQNLTKPGTGISDGDFWIGLLRSGDGQT-SGACPDLYQWSDGSSSQFRNWYTDEPSCGS 141

Qy        119 EVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDE-KPAVPSREAEGEETELT 177
              | ||||||||:|  |:||||::||||||||||:|:|| |  |  |  |:     |:  ||
Db        142 EKCVVMYHQPTANPGLGGPYLYQWNDDRCNMKHNYICTYEPEIHPTEPA-----EKPYLT 196

Qy        178 TPVLPEETQE 187
                   |||| |
Db        197 NQ--PEETHE 204


RESULT 2
CHOD_HUMAN
ID   CHOD_HUMAN     STANDARD;      PRT;   273 AA.
AC   Q9H9P2; Q9HCY3;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   Chondrolectin precursor (Transmembrane protein MT75) (Protein PRED12).
GN   CHODL OR C21ORF68.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.
RX   MEDLINE=22074930; PubMed=12079284;
RA   Weng L., Smits P., Wauters J., Merregaert J.;
RT   "Molecular cloning and characterization of human chondrolectin, a
RT   novel type I transmembrane protein homologous to C-type lectins.";
RL   Genomics 80:62-70(2002).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20289799; PubMed=10830953;
RA   Hattori M., Fujiyama A., Taylor T.D., Watanabe H., Yada T.,
RA   Park H.-S., Toyoda A., Ishii K., Totoki Y., Choi D.-K., Groner Y.,
RA   Soeda E., Ohki M., Takagi T., Sakaki Y., Taudien S., Blechschmidt K.,
RA   Polley A., Menzel U., Delabar J., Kumpf K., Lehmann R., Patterson D.,
RA   Reichwald K., Rump A., Schillhabel M., Schudy A., Zimmermann W.,
RA   Rosenthal A., Kudoh J., Shibuya K., Kawasaki K., Asakawa S.,
RA   Shintani A., Sasaki T., Nagamine K., Mitsuyama S., Antonarakis S.E.,
RA   Minoshima S., Shimizu N., Nordsiek G., Hornischer K., Brandt P.,
RA   Scharfe M., Schoen O., Desario A., Reichelt J., Kauer G., Bloecker H.,
RA   Ramser J., Beck A., Klages S., Hennig S., Riesselmann L., Dagand E.,
RA   Wehrmeyer S., Borzym K., Gardiner K., Nizetic D., Francis F.,
RA   Lehrach H., Reinhardt R., Yaspo M.-L.;
RT   "The DNA sequence of human chromosome 21.";
RL   Nature 405:311-319(2000).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Brain;
RX   MEDLINE=22388257; PubMed=12477932;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length
RT   human and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [4]
RP   SEQUENCE OF 27-273 FROM N.A.
RA   Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
RA   Nishikawa T., Nagai K., Sugano S., Shiratori A., Sudo H.,
RA   Wagatsuma M., Hosoiri T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,
RA   Takahashi M., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,
RA   Watanabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K.,
RA   Yamamoto J., Wakamatsu A., Nakamura Y., Nagahari K., Masuho Y.,
RA   Ninomiya K., Iwayanagi T.;
RT   "NEDO human cDNA sequencing project.";
RL   Submitted (AUG-2000) to the EMBL/GenBank/DDBJ databases.
RN   [5]
RP   TISSUE SPECIFICITY.
RX   MEDLINE=21564202; PubMed=11707072;
RA   Reymond A., Friedli M., Neergard Henrichsen C., Chapot F., Deutsch S.,
RA   Ucla C., Rossier C., Lyle R., Guipponi M., Antonarakis S.E.;
RT   "From PREDs and open reading frames to cDNA isolation: revisiting the
RT   human chromosome 21 transcription map.";
RL   Genomics 78:46-54(2001).
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein (Potential).
CC   -!- TISSUE SPECIFICITY: Found in spleen, testis, prostate and fetal
CC       liver. Expression limited to vascular muscle of testis, smooth
CC       muscle of prostate stroma, heart muscle, skeletal muscle, crypts
CC       of small intestine, and red pulp of spleen.
CC   -!- PTM: N-glycosylated.
CC   -!- SIMILARITY: Contains 1 C-type lectin family domain.
CC   ---------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; AF257472; AAL05981.1; -.
DR   EMBL; AL163217; CAB90388.1; -.
DR   EMBL; BC009418; AAH09418.1; -.
DR   EMBL; AK022689; BAB14181.1; ALT_INIT.
```

```
DR   HSSP; P22897; 1EGG.
DR   Genew; HGNC:17807; CHODL.
DR   MIM; 607247; -.
DR   InterPro; IPR001304; Lectin_C.
DR   Pfam; PF00059; lectin_c; 1.
DR   SMART; SM00034; CLECT; 1.
DR   PROSITE; PS00615; C_TYPE_LECTIN_1; FALSE_NEG.
DR   PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
KW   Lectin; Transmembrane; Glycoprotein; Signal.
FT   SIGNAL        1     21       POTENTIAL.
FT   CHAIN        22    273       CHONDROLECTIN.
FT   DOMAIN       22    216       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    217    237       POTENTIAL.
FT   DOMAIN      238    273       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN       35    179       C-TYPE LECTIN.
FT   CARBOHYD     86     86       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   273 AA;  30431 MW;  F4890AAFB572A311 CRC64;

  Query Match             50.4%;  Score 561.5;  DB 1;  Length 273;
  Best Local Similarity   60.1%;  Pred. No. 3.3e-43;
  Matches  101;  Conservative   25;  Mismatches   35;  Indels    7;  Gaps    3;

Qy          4 RLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEK 63
              |::||| ||    : ||||: |||: | |::|:||: ||  :|| |:|:|:| ||||||
Db         23 RVVSGQKVCFADFKHPCYKMAYFHELSSRVSFQEARLACESEGGVLLSLENEAEQKLIES 82

Qy         64 FIENLLP-----SDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEPSCGS 118
                ::||       |||||||||| |  : |: | || ||| |:||| ||:|||| |||||||
Db         83 MLQNLTKPGTGISDGDFWIGLWRNGDGQT-SGACPDLYQWSDGSNSQYRNWYTDEPSCGS 141

Qy        119 EVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDE-KPAVP 165
              | |||||||||:|  |:||||::||||||||||:|:|||  |  |  |
Db        142 EKCVVMYHQPTANPGLGGPYLYQWNDDRCNMKHNYICKYEPEINPTAP 189


RESULT 3
MANR_HUMAN
ID   MANR_HUMAN     STANDARD;      PRT;  1456 AA.
AC   P22897;
DT   01-AUG-1991 (Rel. 19, Created)
DT   01-AUG-1991 (Rel. 19, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   Macrophage mannose receptor precursor (MMR) (CD206 antigen).
GN   MRC1.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC   TISSUE=Placenta;
RX   MEDLINE=90324192; PubMed=2373685;
RA   Taylor M.E., Conary J.T., Lennartz M.R., Stahl P.D., Drickamer K.;
RT   "Primary structure of the mannose receptor contains multiple motifs
RT   resembling carbohydrate-recognition domains.";
RL   J. Biol. Chem. 265:12156-12162(1990).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=93052405; PubMed=1294118;
RA   Kim S.J., Ruiz N., Bezouska K., Drickamer K.;
RT   "Organization of the gene encoding the human macrophage mannose
RT   receptor (MRC1).";
RL   Genomics 14:721-727(1992).
RN   [3]
RP   STUDIES ON THE BINDING OF INDIVIDUAL LECTIN DOMAINS.
RX   MEDLINE=92112893; PubMed=1730714;
RA   Taylor M.E., Bezouska K., Drickamer K.;
RT   "Contribution to ligand binding by multiple carbohydrate-recognition
RT   domains in the macrophage mannose receptor.";
RL   J. Biol. Chem. 267:1719-1726(1992).
RN   [4]
RP   X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS) OF 642-788.
RX   MEDLINE=20347275; PubMed=10779515;
RA   Feinberg H., Park-Snyder S., Kolatkar A.R., Heise C.T., Taylor M.E.,
RA   Weis W.I.;
RT   "Structure of a C-type carbohydrate recognition domain from the
RT   macrophage mannose receptor.";
RL   J. Biol. Chem. 275:21539-21548(2000).
CC   -!- FUNCTION: Mediates the endocytosis of glycoproteins by
CC       macrophages, in several recognition and uptake processes.
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC   -!- MISCELLANEOUS: CRDs 1-3 have at most very weak affinity for
CC       carbohydrate. CRD 4 shows the highest affinity binding and has
CC       multispecificity for a variety of monosaccharides. At least 3 CRDs
CC       (4, 5, and 7) are required for high affinity binding and
CC       endocytosis of multivalent glycoconjugates.
CC   -!- SIMILARITY: Contains 8 C-type lectin family domains.
CC   -!- SIMILARITY: Contains 1 ricin B-type lectin domain.
CC   -!- DATABASE: NAME=PROW; NOTE=PROW 2:85-89(2001);
CC       WWW="http://www.ncbi.nlm.nih.gov/prow/guide/1644341535_g.htm".
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; J05550; AAA59868.1; -.
DR   EMBL; M93221; AAA60389.1; -.
DR   EMBL; M93192; AAA60389.1; JOINED.
DR   EMBL; M93193; AAA60389.1; JOINED.
DR   EMBL; M93194; AAA60389.1; JOINED.
DR   EMBL; M93195; AAA60389.1; JOINED.
DR   EMBL; M93196; AAA60389.1; JOINED.
DR   EMBL; M93197; AAA60389.1; JOINED.
DR   EMBL; M93198; AAA60389.1; JOINED.
DR   EMBL; M93199; AAA60389.1; JOINED.
DR   EMBL; M93200; AAA60389.1; JOINED.
DR   EMBL; M93201; AAA60389.1; JOINED.
DR   EMBL; M93202; AAA60389.1; JOINED.
DR   EMBL; M93203; AAA60389.1; JOINED.
DR   EMBL; M93204; AAA60389.1; JOINED.
DR   EMBL; M93205; AAA60389.1; JOINED.
DR   EMBL; M93206; AAA60389.1; JOINED.
DR   EMBL; M93207; AAA60389.1; JOINED.
DR   EMBL; M93208; AAA60389.1; JOINED.
DR   EMBL; M93209; AAA60389.1; JOINED.
DR   EMBL; M93210; AAA60389.1; JOINED.
DR   EMBL; M93211; AAA60389.1; JOINED.
DR   EMBL; M93212; AAA60389.1; JOINED.
DR   EMBL; M93213; AAA60389.1; JOINED.
DR   EMBL; M93214; AAA60389.1; JOINED.
DR   EMBL; M93215; AAA60389.1; JOINED.
DR   EMBL; M93216; AAA60389.1; JOINED.
DR   EMBL; M93217; AAA60389.1; JOINED.
DR   EMBL; M93218; AAA60389.1; JOINED.
DR   EMBL; M93219; AAA60389.1; JOINED.
DR   EMBL; M93220; AAA60389.1; JOINED.
DR   PIR; A36563; A36563.
DR   PDB; 1EGG; 30-AUG-00.
DR   PDB; 1EGI; 30-AUG-00.
DR   Genew; HGNC:7228; MRC1.
DR   MIM; 153618; -.
DR   GO; GO:0005887; C:integral to plasma membrane; TAS.
DR   GO; GO:0005537; F:mannose binding; TAS.
DR   GO; GO:0004872; F:receptor activity; TAS.
DR   GO; GO:0006898; P:receptor mediated endocytosis; TAS.
DR   InterPro; IPR002353; AntifreezeII.
DR   InterPro; IPR000562; FN_Type_II.
DR   InterPro; IPR001304; Lectin_C.
DR   InterPro; IPR000772; Ricin_B_lectin.
DR   InterPro; IPR008997; RicinB_like.
DR   Pfam; PF00040; fn2; 1.
DR   Pfam; PF00059; lectin_c; 8.
```

```
DR   Pfam; PF00652; Ricin_B_lectin; 2.
DR   PRINTS; PR00356; ANTIFREEZEII.
DR   PRINTS; PR00013; FNTYPEII.
DR   ProDom; PD000995; FN_Type_II; 1.
DR   SMART; SM00034; CLECT; 8.
DR   SMART; SM00059; FN2; 1.
DR   SMART; SM00458; RICIN; 1.
DR   PROSITE; PS00615; C_TYPE_LECTIN_1; 6.
DR   PROSITE; PS50041; C_TYPE_LECTIN_2; 8.
DR   PROSITE; PS00023; FIBRONECTIN_2; 1.
DR   PROSITE; PS50231; RICIN_B_LECTIN; 1.
KW   Receptor; Signal; Calcium-binding; Transmembrane; Repeat;
KW   Glycoprotein; Endocytosis; Macrophage; Lectin; Antigen; 3D-structure.
FT   SIGNAL        1     18       POTENTIAL.
FT   CHAIN        19   1456       MACROPHAGE MANNOSE RECEPTOR.
FT   DOMAIN       19   1383       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   1384   1411       POTENTIAL.
FT   DOMAIN     1412   1456       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN       22    142       RICIN B-TYPE LECTIN.
FT   DOMAIN      157    212       FIBRONECTIN TYPE-II.
FT   DOMAIN      216    344       C-TYPE LECTIN 1 (LONG FORM).
FT   DOMAIN      360    490       C-TYPE LECTIN 2 (LONG FORM).
FT   DOMAIN      502    629       C-TYPE LECTIN 3 (LONG FORM).
FT   DOMAIN      644    781       C-TYPE LECTIN 4 (LONG FORM).
FT   DOMAIN      805    926       C-TYPE LECTIN 5 (LONG FORM).
FT   DOMAIN      943   1083       C-TYPE LECTIN 6 (LONG FORM).
FT   DOMAIN     1100   1216       C-TYPE LECTIN 7 (LONG FORM).
FT   DOMAIN     1228   1359       C-TYPE LECTIN 8 (LONG FORM).
FT   DISULFID    646    659
FT   DISULFID    680    777
FT   DISULFID    753    769
FT   CARBOHYD    104    104       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    344    344       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    529    529       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    926    926       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    930    930       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1160   1160       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1205   1205       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1311   1311       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   TURN        648    649
FT   STRAND      651    652
FT   TURN        654    655
FT   STRAND      658    663
FT   HELIX       667    669
FT   STRAND      671    671
FT   HELIX       673    683
FT   TURN        684    684
FT   STRAND      686    687
FT   HELIX       693    705
FT   TURN        706    707
FT   TURN        709    710
FT   STRAND      712    718
FT   TURN        723    724
FT   STRAND      727    727
FT   TURN        729    730
FT   STRAND      733    733
FT   TURN        741    742
FT   HELIX       746    748
FT   STRAND      752    757
FT   TURN        758    761
FT   STRAND      764    768
FT   TURN        769    770
FT   STRAND      773    780
FT   TURN        781    782
SQ   SEQUENCE   1456 AA;  166011 MW;  264E5AF3C576A5E3 CRC64;

  Query Match             16.3%;  Score 182;  DB 1;  Length 1456;
  Best Local Similarity   25.5%;  Pred. No. 3.2e-08;
  Matches   52;  Conservative   37;  Mismatches   61;  Indels   54;  Gaps    8;

Qy         21 YKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFWIGLR 80
              ||   :: :  :   : |:  |:|: | ||||:|| |:| : |:: |   :   ::|||
Db        807 YKDYQYYFSKEKETMDNARAFCKRNFGDLVSIQSESEKKFLWKYV-NRNDAQSAYFIGLL 865

Qy         81 RREEKQSNSTACQDLYAWTDGSISQFRNWYVDEPSCGS--EVCVVMYHQPSAPAGIGGPY 138
                 :|:          :|| |||   : :|   ||:  :  | || ||    : :|
Db        866 ISLDKK---------FAWMDGSKVDYVSWATGEPNFANEDENCVTMY----SNSGF---- 908

Qy        139 MFQWNDDRCNMKNNFICKYSDEK----PAVPSREAEGEETELTTPVLPEETQE------- 187
                |||  |   | |||:  :         :|            | | :|   :|
Db        909 ---WNDINCGYPNAFICQRHNSSINATTVMP----------TMPSVPSGCKEGWNFYSN 954

Qy        188 ---------EDAKKTFKESREAAL 202
                      |: :| ::|:|:| :
Db        955 KCFKIFGFMEEERKNWQEARKACI 978


RESULT 4
PGCN_MOUSE
ID   PGCN_MOUSE     STANDARD;      PRT;  1268 AA.
AC   P55066;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Neurocan core protein precursor (Chondroitin sulfate proteoglycan 3).
GN   CSPG3 OR NCAN.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BALB/c; TISSUE=Brain;
RX   MEDLINE=96039250; PubMed=7490074;
RA   Rauch U., Grimpe B., Kulbe G., Arnold-Ammer I., Beier D.,
RA   Faessler R.;
RT   "Structure and chromosomal localization of the mouse neurocan gene.";
RL   Genomics 28:405-410(1995).
CC   -!- FUNCTION: May modulate neuronal adhesion and neurite growth during
CC       development by binding to neural cell adhesion molecules (NG-CAM
CC       and N-CAM). Chondroitin sulfate proteoglycan; binds to hyaluronic
CC       acid.
CC   -!- TISSUE SPECIFICITY: Brain.
CC   -!- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
CC   -!- SIMILARITY: Contains 2 EGF-like domains.
CC   -!- SIMILARITY: Contains 2 link domains.
CC   -!- SIMILARITY: Contains 1 C-type lectin family domain.
CC   -!- SIMILARITY: Contains 1 Sushi (SCR) domain.
CC   -!- SIMILARITY: Belongs to the aggrecan/versican proteoglycan family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X84727; CAA59216.1; -.
DR   PIR; S52781; S52781.
DR   HSSP; P00740; 1EDM.
DR   MGD; MGI:104694; Cspg3.
DR   InterPro; IPR002353; AntifreezeII.
DR   InterPro; IPR000152; Asx_hydroxyl_S.
DR   InterPro; IPR000742; EGF_2.
DR   InterPro; IPR001881; EGF_Ca.
DR   InterPro; IPR006209; EGF_like.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003599; Ig.
DR   InterPro; IPR001304; Lectin_C.
DR   InterPro; IPR000538; Link.
DR   InterPro; IPR000436; Sushi_SCR_CCP.
DR   Pfam; PF00008; EGF; 2.
DR   Pfam; PF00047; ig; 1.
DR   Pfam; PF00059; lectin_c; 1.
DR   Pfam; PF00084; sushi; 1.
```

```
DR   Pfam; PF00193; Xlink; 2.
DR   PRINTS; PR00356; ANTIFREEZEII.
DR   PRINTS; PR01265; LINKMODULE.
DR   ProDom; PD000918; Link; 2.
DR   SMART; SM00032; CCP; 1.
DR   SMART; SM00034; CLECT; 1.
DR   SMART; SM00179; EGF_CA; 1.
DR   SMART; SM00409; IG; 1.
DR   SMART; SM00445; LINK; 2.
DR   PROSITE; PS00010; ASX_HYDROXYL; 1.
DR   PROSITE; PS00615; C_TYPE_LECTIN_1; 1.
DR   PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
DR   PROSITE; PS00022; EGF_1; 3.
DR   PROSITE; PS01186; EGF_2; 1.
DR   PROSITE; PS50026; EGF_3; 2.
DR   PROSITE; PS01187; EGF_CA; 1.
DR   PROSITE; PS50835; IG_LIKE; 1.
DR   PROSITE; PS01241; LINK; 2.
KW   Glycoprotein; Hyaluronic acid; Proteoglycan; Immunoglobulin domain;
KW   EGF-like domain; Calcium; Repeat; Lectin; Sushi; Signal.
FT   SIGNAL        1     22       POTENTIAL.
FT   CHAIN        23   1268       NEUROCAN CORE PROTEIN.
FT   DOMAIN       37    157       IG-LIKE V-TYPE.
FT   DOMAIN      158    253       LINK 1.
FT   DOMAIN      259    355       LINK 2.
FT   DOMAIN      960    996       EGF-LIKE 1.
FT   DOMAIN      998   1034       EGF-LIKE 2, CALCIUM-BINDING (POTENTIAL).
FT   DOMAIN     1036   1165       C-TYPE LECTIN.
FT   DOMAIN     1166   1224       SUSHI.
FT   DISULFID     58    139       BY SIMILARITY.
FT   DISULFID    181    252       BY SIMILARITY.
FT   DISULFID    205    226       BY SIMILARITY.
FT   DISULFID    279    354       BY SIMILARITY.
FT   DISULFID    303    324       BY SIMILARITY.
FT   DISULFID    964    975       BY SIMILARITY.
FT   DISULFID    969    984       BY SIMILARITY.
FT   DISULFID    986    995       BY SIMILARITY.
FT   DISULFID   1040   1051       BY SIMILARITY.
FT   DISULFID   1068   1160       BY SIMILARITY.
FT   DISULFID   1136   1152       BY SIMILARITY.
FT   DISULFID   1167   1210       BY SIMILARITY.
FT   DISULFID   1196   1223       BY SIMILARITY.
FT   CARBOHYD    121    121       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    339    339       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    742    742       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    978    978       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1175   1175       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   1268 AA;  137200 MW;  3014E8E202A2FAEC CRC64;

  Query Match             15.9%;  Score 177;  DB 1;  Length 1268;
  Best Local Similarity   31.5%;  Pred. No. 7.6e-08;
  Matches   45;  Conservative   17;  Mismatches   49;  Indels   32;  Gaps    6;

Qy         17 QRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFW 76
              |  ||:  ||   : |  :|:|:  |||  | | |: | :| | |  |         : |
Db       1048 QGHCYR--YF---AHRRAWEDAERDCRRRAGHLTSVHSPEEHKFINSF------GHENSW 1096

Qy         77 IGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEPS---CGSEVCVVMYHQPSAPAG 133
              |||  |  ::         : ||| :  |: ||   :|     | | ||||    |
Db       1097 IGLNDRTVERD--------FQWTDNTGLQYENWREKQPDNFFAGGEDCVVMVAHESG--- 1145

Qy        134 IGGPYMFQWNDDRCNMKNNFICK 156
                :|||  ||    ::||
Db       1146 -------RWNDVPCNYNLPYVCK 1161


RESULT 5
PGCN_HUMAN
ID   PGCN_HUMAN     STANDARD;      PRT;  1321 AA.
AC   O14594; Q9UPK6;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Neurocan core protein precursor (Chondroitin sulfate proteoglycan 3).
GN   CSPG3 OR NCAN OR NEUR.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=99013874; PubMed=9795216;
RA   Prange C.K., Pennacchio L.A., Lieuallen K., Fan W., Lennon G.G.;
RT   "Characterization of the human neurocan gene, CSPG3.";
RL   Gene 221:199-205(1998).
RN   [2]
RP   SEQUENCE OF 1-990 AND 1007-1321 FROM N.A., AND VARIANT ALA-1254.
RA   Lamerdin J.E., McCready P.M., Skowronski E., Adamson A.W.,
RA   Burkhart-Schultz K., Gordon L., Kyle A., Ramirez M., Stilwagen S.,
RA   Phan H., Velasco N., Do L., Regala W., Terry A., Garnes J.,
RA   Danganan L., Poundstone P., Christensen M., Georgescu A., Avila J.,
RA   Liu S., Attix C., Andreise T., Trankheim M., Amico-Keller G.,
RA   Coefield J., Duarte S., Lucas S., Bruce R., Thomas P., Quan G.,
RA   Kronmiller B., Arellano A., Montgomery M., Ow D., Nolan M., Trong S.,
RA   Kobayashi A., Olsen A.S., Carrano A.V.;
RT   "Sequence analysis of an ~1 Mb region containing the MEF2B gene in
RT   19p12.";
RL   Submitted (JUL-1998) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: May modulate neuronal adhesion and neurite growth during
CC       development by binding to neural cell adhesion molecules (NG-CAM
CC       and N-CAM). Chondroitin sulfate proteoglycan; binds to hyaluronic
CC       acid.
CC   -!- TISSUE SPECIFICITY: Brain.
CC   -!- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
CC   -!- SIMILARITY: Contains 2 EGF-like domains.
CC   -!- SIMILARITY: Contains 2 link domains.
CC   -!- SIMILARITY: Contains 1 C-type lectin family domain.
CC   -!- SIMILARITY: Contains 1 Sushi (SCR) domain.
CC   -!- SIMILARITY: Belongs to the aggrecan/versican proteoglycan family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF026547; AAC80576.1; -.
DR   EMBL; AC003110; AAB86655.1; -.
DR   EMBL; AC005254; AAC25581.1; -.
DR   HSSP; P00740; 1EDM.
DR   Genew; HGNC:2465; CSPG3.
DR   MIM; 600826; -.
DR   InterPro; IPR000152; Asx_hydroxyl_S.
DR   InterPro; IPR000742; EGF_2.
DR   InterPro; IPR001881; EGF_Ca.
DR   InterPro; IPR006209; EGF_like.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003599; Ig.
DR   InterPro; IPR001304; Lectin_C.
DR   InterPro; IPR000538; Link.
DR   InterPro; IPR000436; Sushi_SCR_CCP.
DR   Pfam; PF00008; EGF; 2.
DR   Pfam; PF00047; ig; 1.
DR   Pfam; PF00059; lectin_c; 1.
DR   Pfam; PF00084; sushi; 1.
DR   Pfam; PF00193; Xlink; 2.
DR   PRINTS; PR01265; LINKMODULE.
DR   ProDom; PD000918; Link; 2.
DR   SMART; SM00032; CCP; 1.
DR   SMART; SM00034; CLECT; 1.
DR   SMART; SM00179; EGF_CA; 1.
DR   SMART; SM00409; IG; 1.
DR   SMART; SM00445; LINK; 2.
DR   PROSITE; PS00010; ASX_HYDROXYL; 1.
```

```
DR   PROSITE; PS00615; C_TYPE_LECTIN_1; 1.
DR   PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
DR   PROSITE; PS00022; EGF_1; 3.
DR   PROSITE; PS01186; EGF_2; 1.
DR   PROSITE; PS50026; EGF_3; 2.
DR   PROSITE; PS01187; EGF_CA; 1.
DR   PROSITE; PS50835; IG_LIKE; 1.
DR   PROSITE; PS01241; LINK; 2.
KW   Glycoprotein; Hyaluronic acid; Proteoglycan; Immunoglobulin domain;
KW   EGF-like domain; Calcium; Repeat; Lectin; Sushi; Signal; Polymorphism.
FT   SIGNAL       1     22       POTENTIAL.
FT   CHAIN       23   1321       NEUROCAN CORE PROTEIN.
FT   DOMAIN      38    153       IG-LIKE V-TYPE.
FT   DOMAIN     159    254       LINK 1.
FT   DOMAIN     260    356       LINK 2.
FT   DOMAIN    1008   1044       EGF-LIKE 1.
FT   DOMAIN    1046   1082       EGF-LIKE 2, CALCIUM-BINDING (POTENTIAL).
FT   DOMAIN    1084   1213       C-TYPE LECTIN.
FT   DOMAIN    1214   1272       SUSHI.
FT   DISULFID    59    140       BY SIMILARITY.
FT   DISULFID   182    253       BY SIMILARITY.
FT   DISULFID   206    227       BY SIMILARITY.
FT   DISULFID   280    355       BY SIMILARITY.
FT   DISULFID   304    325       BY SIMILARITY.
FT   DISULFID  1012   1023       BY SIMILARITY.
FT   DISULFID  1017   1032       BY SIMILARITY.
FT   DISULFID  1034   1043       BY SIMILARITY.
FT   DISULFID  1088   1099       BY SIMILARITY.
FT   DISULFID  1116   1208       BY SIMILARITY.
FT   DISULFID  1184   1200       BY SIMILARITY.
FT   DISULFID  1215   1258       BY SIMILARITY.
FT   DISULFID  1244   1271       BY SIMILARITY.
FT   CARBOHYD   122    122       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   340    340       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD  1026   1026       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD  1223   1223       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   VARIANT   1254   1254       V -> A (in dbSNP:1064389).
FT                               /FTId=VAR_016176.
FT   CONFLICT  1234   1234       N -> Y (IN REF. 2).
FT   CONFLICT  1282   1282       G -> R (IN REF. 2).
SQ   SEQUENCE  1321 AA;  142972 MW;  2EF47F823DB980B8 CRC64;

  Query Match             15.8%;  Score 176;  DB 1;  Length 1321;
  Best Local Similarity   31.5%;  Pred. No. 9.9e-08;
  Matches   45;  Conservative   18;  Mismatches   48;  Indels   32;  Gaps    6;

Qy         17 QRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFW 76
              |  ||:  ||   : |  :|:|:: |||  | | |: | :|   |  |         : |
Db       1096 QGHCYR--YF---AHRRAWEDAEKDCRRRSGHLTSVHSPEEHSFINSF------GHENTW 1144

Qy         77 IGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEPS---CGSEVCVVMYHQPSAPAG 133
              |||  |  ::         : ||| :  || ||  ::|     | | ||||     |
Db       1145 IGLNDRIVERD--------FQWTDNTGLQFENWRENQPDNFFAGGEDCVVMVAHESG--- 1193

Qy        134 IGGPYMFQWNDDRCNMKNNFICK 156
                :|||  ||    ::||
Db       1194 -------RWNDVPCNYNLPYVCK 1209


RESULT 6
PGCV_BOVIN
ID   PGCV_BOVIN     STANDARD;      PRT;  3381 AA.
AC   P81282; O77609; O77610; O77611; O77612;
DT   15-DEC-1998 (Rel. 37, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Versican core protein precursor (Large fibroblast proteoglycan)
DE   (Chondroitin sulfate proteoglycan core protein 2) (PG-M) (Glial
DE   hyaluronate-binding protein) (GHAP).
GN   CPSG2.
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORMS VO; V1; V2 AND V3).
RC   TISSUE=Forebrain;
RX   MEDLINE=98288320; PubMed=9624174;
RA   Schmalfeldt M., Dours-Zimmermann M.T., Winterhalter K.H.,
RA   Zimmermann D.R.;
RT   "Versican V2 is a major extracellular matrix component of the mature
RT   bovine brain.";
RL   J. Biol. Chem. 273:15758-15764(1998).
RN   [2]
RP   SEQUENCE OF 21-53; 78-96; 226-250; 262-277; 295-306; 314-324; 329-331
RP   AND 342-348.
RC   TISSUE=Spinal cord;
RX   MEDLINE=92062692; PubMed=1720020;
RA   Perides G., Biviano F., Bignami A.;
RT   "Interaction of a brain extracellular matrix protein with hyaluronic
RT   acid.";
RL   Biochim. Biophys. Acta 1075:248-258(1991).
CC   -!- FUNCTION: May play a role in intercellular signaling and in
CC       connecting cells with the extracellular matrix. May take part in
CC       the regulation of cell motility, growth and differentiation. Binds
CC       hyaluronic acid.
CC   -!- SUBUNIT: Interacts with FBLN1 (By similarity).
CC   -!- SUBCELLULAR LOCATION: Secreted; extracellular matrix.
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=4;
CC         Comment=Additional isoforms seem to exist;
CC       Name=V0;
CC         IsoId=P81282-1; Sequence=Displayed;
CC       Name=V1;
CC         IsoId=P81282-2; Sequence=VSP_003078, VSP_003079;
CC       Name=V2;
CC         IsoId=P81282-3; Sequence=VSP_003080;
CC       Name=V3;
CC         IsoId=P81282-4; Sequence=VSP_003078, VSP_003081;
CC   -!- TISSUE SPECIFICITY: Cerebral white matter. V0 and V1 are expressed
CC       in the central nervous system, and in a number of mesenchymal and
CC       epithelial tissues; the major isoform v2 is restricted to the
CC       central nervous system.
CC   -!- DEVELOPMENTAL STAGE: Disappears after the cartilage development
CC       (By similarity).
CC   -!- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
CC   -!- SIMILARITY: Contains 2 link domains.
CC   -!- SIMILARITY: Contains 2 EGF-like domains.
CC   -!- SIMILARITY: Contains 1 C-type lectin family domain.
CC   -!- SIMILARITY: Contains 1 Sushi (SCR) domain.
CC   -!- SIMILARITY: Belongs to the aggrecan/versican proteoglycan family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF060456; AAC24358.1; -.
DR   EMBL; AF060457; AAC24359.1; -.
DR   EMBL; AF060458; AAC24360.1; -.
DR   EMBL; AF060459; AAC24361.1; -.
DR   PIR; T14274; T14274.
DR   PIR; T42389; T42389.
DR   HSSP; P01132; 1EPG.
DR   InterPro; IPR000152; Asx_hydroxyl_S.
DR   InterPro; IPR000742; EGF_2.
DR   InterPro; IPR001881; EGF_Ca.
DR   InterPro; IPR006209; EGF_like.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003599; Ig.
DR   InterPro; IPR001304; Lectin_C.
DR   InterPro; IPR000538; Link.
```

```
DR   InterPro; IPR000436; Sushi_SCR_CCP.
DR   Pfam; PF00008; EGF; 2.
DR   Pfam; PF00047; ig; 1.
DR   Pfam; PF00059; lectin_c; 1.
DR   Pfam; PF00084; sushi; 1.
DR   Pfam; PF00193; Xlink; 2.
DR   PRINTS; PR01265; LINKMODULE.
DR   ProDom; PD000918; Link; 2.
DR   SMART; SM00032; CCP; 1.
DR   SMART; SM00034; CLECT; 1.
DR   SMART; SM00179; EGF_CA; 1.
DR   SMART; SM00409; IG; 1.
DR   SMART; SM00445; LINK; 2.
DR   PROSITE; PS00010; ASX_HYDROXYL; 1.
DR   PROSITE; PS00615; C_TYPE_LECTIN_1; 1.
DR   PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
DR   PROSITE; PS00022; EGF_1; 2.
DR   PROSITE; PS01186; EGF_2; 1.
DR   PROSITE; PS50026; EGF_3; 2.
DR   PROSITE; PS01187; EGF_CA; 1.
DR   PROSITE; PS50835; IG_LIKE; 1.
DR   PROSITE; PS01241; LINK; 2.
KW   Glycoprotein; Proteoglycan; Lectin; Extracellular matrix; Sushi;
KW   Signal; Repeat; EGF-like domain; Calcium; Immunoglobulin domain;
KW   Hyaluronic acid; Alternative splicing.
FT   SIGNAL        1     20       POTENTIAL.
FT   CHAIN        21   3381       VERSICAN CORE PROTEIN.
FT   DOMAIN       21    147       IG-LIKE V-TYPE.
FT   DOMAIN      168    245       LINK 1.
FT   DOMAIN      266    347       LINK 2.
FT   DOMAIN      349   1336       GAG-ALPHA
FT                                (GLUCOSAMINOGLYCAN ATTACHMENT DOMAIN).
FT   DOMAIN     1337   3074       GAG-BETA.
FT   DOMAIN     3074   3110       EGF-LIKE 1.
FT   DOMAIN     3112   3148       EGF-LIKE 2, CALCIUM-BINDING.
FT   DOMAIN     3161   3275       C-TYPE LECTIN.
FT   DOMAIN     3280   3338       SUSHI.
FT   DISULFID     44    131       BY SIMILARITY.
FT   DISULFID    173    244       BY SIMILARITY.
FT   DISULFID    197    218       BY SIMILARITY.
FT   DISULFID    271    346       BY SIMILARITY.
FT   DISULFID    295    316       BY SIMILARITY.
FT   DISULFID   3078   3089       BY SIMILARITY.
FT   DISULFID   3083   3098       BY SIMILARITY.
FT   DISULFID   3100   3109       BY SIMILARITY.
FT   DISULFID   3116   3127       BY SIMILARITY.
FT   DISULFID   3121   3136       BY SIMILARITY.
FT   DISULFID   3138   3147       BY SIMILARITY.
FT   DISULFID   3154   3165       BY SIMILARITY.
FT   DISULFID   3182   3274       BY SIMILARITY.
FT   DISULFID   3250   3266       BY SIMILARITY.
FT   DISULFID   3281   3324       BY SIMILARITY.
FT   DISULFID   3310   3337       BY SIMILARITY.
FT   CARBOHYD     57     57       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    331    331       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    352    352       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    817    817       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    965    965       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1017   1017       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1333   1333       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1393   1393       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1437   1437       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1463   1463       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1653   1653       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1974   1974       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   2045   2045       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   2074   2074       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   2103   2103       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   2263   2263       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   2290   2290       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   2356   2356       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   2623   2623       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   2641   2641       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   2919   2919       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   3052   3052       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   3354   3354       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   3364   3364       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   VARSPLIC    349    349       P -> R (in isoform V1 and isoform V3).
FT                                /FTId=VSP_003078.
FT   VARSPLIC    350   1336       Missing (in isoform V1).
FT                                /FTId=VSP_003079.
FT   VARSPLIC   1337   3074       Missing (in isoform V2).
FT                                /FTId=VSP_003080.
FT   VARSPLIC    350   3074       Missing (in isoform V3).
FT                                /FTId=VSP_003081.
FT   CONFLICT     25     25       MISSING (IN REF. 2).
FT   CONFLICT     51     51       MISSING (IN REF. 2).
FT   CONFLICT     89     89       N -> D (IN REF. 2).
FT   CONFLICT     96     96       Q -> D (IN REF. 2).
FT   CONFLICT    346    346       C -> R (IN REF. 2).
SQ   SEQUENCE   3381 AA;  369984 MW;  F09716FA7778D459 CRC64;
```

Query Match 15.7%; Score 174.5; DB 1; Length 3381;
Best Local Similarity 25.5%; Pred. No. 4.3e-07;
Matches 59; Conservative 26; Mismatches 71; Indels 75; Gaps 11;

```
Qy        17 QRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDF- 75
             |  |||  ||   : |  :: |:  ||  |  | || | :||  : :          |:
Db      3162 QGQCYK--YF---AHRRTWDAAERECRLQGAHLTSILSHEEQMFVNRV-------GHDYQ 3209

Qy        76 WIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEP----SCGSEVCVVMYHQPSAP 131
             ||||         |    :  : |||||  |: ||  ::|    | | :  |:::|:
Db      3210 WIGL--------NDKMFEHDFRWTDGSTLQYENWRPNQPDSFFSTGEDCVVIIWHENG-- 3259

Qy       132 AGIGGPYMFQWNDDRCNMKNNFICKYS----DEKPAVPSREAEGE-------------- 172
                 ||||  ||    : ||      : | | : :  |:
Db      3260 ---------QWNDVPCNYHLTYTCKKGTVACGQPPVVENAKTFGKMKPRYEINSLIRYHC 3310

Qy       173 -----ETELTT----------PVL----PEETQEEDAKKTFKESREAALN 203
                 :  | |           | :    |   |   :|| || |  |  |
Db      3311 KDGFIQRHLPTIRCLGNGRWAMPKITCLNPSAYQRTYSKKYFKNSSSAKDN 3361
```

```
RESULT 7
PGCN_RAT
ID   PGCN_RAT       STANDARD;      PRT;  1257 AA.
AC   P55067;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   Neurocan core protein precursor (Chondroitin sulfate proteoglycan 3)
DE   (245 kDa early postnatal core glycoprotein) [Contains: 150 kDa adult
DE   core glycoprotein].
GN   CSPG3 OR NCAN.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC   STRAIN=Sprague-Dawley; TISSUE=Brain;
RX   MEDLINE=92406907; PubMed=1326557;
RA   Rauch U., Karthikeyan L., Maurel P., Margolis R.U., Margolis R.K.;
RT   "Cloning and primary structure of neurocan, a developmentally
RT   regulated, aggregating chondroitin sulfate proteoglycan of brain.";
RL   J. Biol. Chem. 267:19536-19547(1992).
RN   [2]
RP   CHARACTERIZATION.
RX   MEDLINE=94230574; PubMed=7513709;
RA   Friedlander D.R., Milev P., Karthikeyan L., Margolis R.K.,
RA   Margolis R.U., Grumet M.;
RT   "The neuronal chondroitin sulfate proteoglycan neurocan binds to the
RT   neural cell adhesion molecules Ng-CAM/L1/NILE and N-CAM, and inhibits
RT   neuronal adhesion and neurite outgrowth.";
RL   J. Cell Biol. 125:669-680(1994).
CC   -!- FUNCTION: May modulate neuronal adhesion and neurite growth during
```

```
CC       development by binding to neural cell adhesion molecules (NG-CAM
CC       and N-CAM). Chondroitin sulfate proteoglycan; binds to hyaluronic
CC       acid.
CC   -!- TISSUE SPECIFICITY: Early postnatal and adult brain; not expressed
CC       in kidney, lung, liver and muscle.
CC   -!- PTM: CONTAINS MOSTLY CHONDROITIN SULFATE, BUT ALSO N-LINKED AND
CC       O-LINKED OLIGOSACCHARIDES (BY SIMILARITY).
CC   -!- PTM: TWO ISOFORMS WERE FOUND THAT PROBABLY ARISE BY PROTEOLYTIC
CC       DEGRADATION. THE LARGE ISOFORM IS PREDOMINANT IN EARLY POSTNATAL
CC       BRAIN, THE SMALL ISOFORM IS FOUND IN ADULT BRAIN.
CC   -!- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
CC   -!- SIMILARITY: Contains 2 EGF-like domains.
CC   -!- SIMILARITY: Contains 2 link domains.
CC   -!- SIMILARITY: Contains 1 C-type lectin family domain.
CC   -!- SIMILARITY: Contains 1 Sushi (SCR) domain.
CC   -!- SIMILARITY: Belongs to the aggrecan/versican proteoglycan family.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; M97161; AAC37679.1; -.
DR   PIR; S28764; S28764.
DR   HSSP; P00740; 1EDM.
DR   InterPro; IPR002353; AntifreezeII.
DR   InterPro; IPR000152; Asx_hydroxyl_S.
DR   InterPro; IPR000742; EGF_2.
DR   InterPro; IPR001881; EGF_Ca.
DR   InterPro; IPR006209; EGF_like.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003599; Ig.
DR   InterPro; IPR001304; Lectin_C.
DR   InterPro; IPR000538; Link.
DR   InterPro; IPR000436; Sushi_SCR_CCP.
DR   Pfam; PF00008; EGF; 2.
DR   Pfam; PF00047; ig; 1.
DR   Pfam; PF00059; lectin_c; 1.
DR   Pfam; PF00084; sushi; 1.
DR   Pfam; PF00193; Xlink; 2.
DR   PRINTS; PR00356; ANTIFREEZEII.
DR   PRINTS; PR01265; LINKMODULE.
DR   ProDom; PD000918; Link; 2.
DR   SMART; SM00032; CCP; 1.
DR   SMART; SM00034; CLECT; 1.
DR   SMART; SM00179; EGF_CA; 1.
DR   SMART; SM00409; IG; 1.
DR   SMART; SM00445; LINK; 2.
DR   PROSITE; PS00010; ASX_HYDROXYL; 1.
DR   PROSITE; PS00615; C_TYPE_LECTIN_1; 1.
DR   PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
DR   PROSITE; PS00022; EGF_1; 3.
DR   PROSITE; PS01186; EGF_2; 1.
DR   PROSITE; PS50026; EGF_3; 2.
DR   PROSITE; PS01187; EGF_CA; 1.
DR   PROSITE; PS50835; IG_LIKE; 1.
DR   PROSITE; PS01241; LINK; 2.
KW   Glycoprotein; Hyaluronic acid; Proteoglycan; Immunoglobulin domain;
KW   EGF-like domain; Calcium; Repeat; Lectin; Sushi; Signal.
FT   SIGNAL       1     22
FT   CHAIN       23   1257       NEUROCAN CORE PROTEIN.
FT   CHAIN      639   1257       150 kDa ADULT CORE GLYCOPROTEIN.
FT   DOMAIN      37    157       IG-LIKE V-TYPE.
FT   DOMAIN     158    253       LINK 1.
FT   DOMAIN     259    355       LINK 2.
FT   DOMAIN     949    985       EGF-LIKE 1.
FT   DOMAIN     987   1023       EGF-LIKE 2, CALCIUM-BINDING (POTENTIAL).
FT   DOMAIN    1025   1154       C-TYPE LECTIN.
FT   DOMAIN    1155   1213       SUSHI.
FT   DISULFID    58    139       BY SIMILARITY.
FT   DISULFID   181    252       BY SIMILARITY.
FT   DISULFID   205    226       BY SIMILARITY.
FT   DISULFID   279    354       BY SIMILARITY.
FT   DISULFID   303    324       BY SIMILARITY.
FT   DISULFID   953    964       BY SIMILARITY.
FT   DISULFID   958    973       BY SIMILARITY.
FT   DISULFID   975    984       BY SIMILARITY.
FT   DISULFID  1029   1040       BY SIMILARITY.
FT   DISULFID  1057   1149       BY SIMILARITY.
FT   DISULFID  1125   1141       BY SIMILARITY.
FT   DISULFID  1156   1199       BY SIMILARITY.
FT   DISULFID  1185   1212       BY SIMILARITY.
FT   CARBOHYD   121    121       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   339    339       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   737    737       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   944    944       O-LINKED (XYL. . .) (CHONDROITIN SULFATE).
FT   CARBOHYD   967    967       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD  1164   1164       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   1257 AA;  135544 MW;  992B33DCFA19EE1B CRC64;

  Query Match          15.6%;  Score 174;  DB 1;  Length 1257;
  Best Local Similarity  30.8%;  Pred. No. 1.4e-07;
  Matches   44;  Conservative   18;  Mismatches   49;  Indels   32;  Gaps    6;

Qy         17 QRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFW 76
              |  ||:  ||   : |  :|:|:  |||  | | |: | :| | |  |        : |
Db       1037 QGHCYR--YF---AHRRAWEDAERDCRRRAGHLTSVHSPEEHKFINSF------GHENSW 1085

Qy         77 IGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEPS---CGSEVCVVMYHQPSAPAG 133
              |||  |  ::         : ||| :  |: ||   :|     | | ||||     :
Db       1086 IGLNDRTVERD--------FQWTDNTGLQYENWREKQPDNFFAGGEDCVVMVAHENG--- 1134

Qy        134 IGGPYMFQWNDDRCNMKNNFICK 156
                  :|||  ||    ::||
Db       1135 -------RWNDVPCNYNLPYVCK 1150


RESULT 8
PGCV_RAT
ID   PGCV_RAT       STANDARD;      PRT;  2738 AA.
AC   Q9ERB4; O08592; O88564; Q9R1K4;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Versican core protein precursor (Large fibroblast proteoglycan)
DE   (Chondroitin sulfate proteoglycan core protein 2) (PG-M) (Glial
DE   hyaluronate-binding protein) (GHAP) (Fragments).
GN   CPSG2.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE OF 349-2738 FROM N.A. (ISOFORM V0), SEQUENCE FROM N.A.
RP   (ISOFORM V3), AND SEQUENCE OF 2657-2738 FROM N.A. (ISOFORM VINT).
RC   STRAIN=Wistar Kyoto;
RX   MEDLINE=99327053; PubMed=10397680;
RA   Lemire J.M., Braun K.R., Maurel P., Kaplan E.D., Schwartz S.M.,
RA   Wight T.N.;
RT   "Versican/PG-M isoforms in vascular smooth muscle cells.";
RL   Arterioscler. Thromb. Vasc. Biol. 19:1630-1639(1999).
RN   [2]
RP   SEQUENCE OF 349-2738 FROM N.A. (ISOFORM V0).
RC   STRAIN=Wistar Kyoto;
RX   MEDLINE=98308094; PubMed=9642104;
RA   Milev P., Maurel P., Chiba A., Mevissen M., Popp S., Yamaguchi Y.,
RA   Margolis R.K., Margolis R.U.;
RT   "Differential regulation of expression of hyaluronan-binding
RT   proteoglycans in developing brain: aggrecan, versican, neurocan, and
RT   brevican.";
RL   Biochem. Biophys. Res. Commun. 247:207-212(1998).
RN   [3]
RP   SEQUENCE OF 2421-2463 FROM N.A. (ISOFORM V0).
```

```
RC   TISSUE=Kidney;
RX   MEDLINE=98094159; PubMed=9434070;
RA   Pyke C., Kristensen P., Ostergaard P.B., Oturai P.S., Romer J.;
RT   "Proteoglycan expression in the normal rat kidney.";
RL   Nephron 77:461-470(1997).
RN   [4]
RP   SEQUENCE OF 2535-2738 FROM N.A.
RC   STRAIN=Sprague-Dawley; TISSUE=Lung;
RA   Blomberg L.A., Chan W.-Y., Clerch L., Massaro D.;
RT   "Molecular cloning and characterization of two developmentally
RT   regulated genes in rat lung.";
RL   Submitted (SEP-2000) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: May play a role in intercellular signaling and in
CC       connecting cells with the extracellular matrix. May take part in
CC       the regulation of cell motility, growth and differentiation. Binds
CC       hyaluronic acid.
CC   -!- SUBUNIT: Interacts with FBLN1 (By similarity).
CC   -!- SUBCELLULAR LOCATION: Secreted; extracellular matrix.
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=3;
CC         Comment=Additional isoforms seem to exist;
CC       Name=V0;
CC         IsoId=Q9ERB4-1; Sequence=Displayed;
CC       Name=V3;
CC         IsoId=Q9ERB4-2; Sequence=VSP_003091;
CC       Name=Vint;
CC         IsoId=Q9ERB4-3; Sequence=VSP_003092;
CC   -!- TISSUE SPECIFICITY: In kidney is expressed in the papillary area,
CC       but not in glomeruli.
CC   -!- DEVELOPMENTAL STAGE: Disappears after the cartilage development
CC       (By similarity).
CC   -!- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
CC   -!- SIMILARITY: Contains 2 link domains.
CC   -!- SIMILARITY: Contains 2 EGF-like domains.
CC   -!- SIMILARITY: Contains 1 C-type lectin family domain.
CC   -!- SIMILARITY: Contains 1 Sushi (SCR) domain.
CC   -!- SIMILARITY: Belongs to the aggrecan/versican proteoglycan family.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; AF062402; AAC40166.1; -.
DR   EMBL; U75306; AAB51125.1; -.
DR   EMBL; AF084544; AAD48544.1; -.
DR   EMBL; AF072892; AAC26116.1; -.
DR   EMBL; AY007691; AAG16631.1; -.
DR   HSSP; P01132; 1EPG.
DR   InterPro; IPR000152; Asx_hydroxyl_S.
DR   InterPro; IPR000742; EGF_2.
DR   InterPro; IPR001881; EGF_Ca.
DR   InterPro; IPR006209; EGF_like.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003599; Ig.
DR   InterPro; IPR001304; Lectin_C.
DR   InterPro; IPR000538; Link.
DR   InterPro; IPR000436; Sushi_SCR_CCP.
DR   PRINTS; PR01265; LINKMODULE.
DR   ProDom; PD000918; Link; 2.
DR   SMART; SM00032; CCP; 1.
DR   SMART; SM00034; CLECT; 1.
DR   SMART; SM00179; EGF_CA; 1.
DR   SMART; SM00409; IG; 1.
DR   SMART; SM00445; LINK; 2.
DR   PROSITE; PS00010; ASX_HYDROXYL; 1.
DR   PROSITE; PS00615; C_TYPE_LECTIN_1; 1.
DR   PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
DR   PROSITE; PS00022; EGF_1; 2.
DR   PROSITE; PS01186; EGF_2; 1.
DR   PROSITE; PS50026; EGF_3; 2.
DR   PROSITE; PS01187; EGF_CA; 1.
DR   PROSITE; PS50835; IG_LIKE; 1.
DR   PROSITE; PS01241; LINK; 2.
KW   Glycoprotein; Proteoglycan; Lectin; Extracellular matrix; Sushi;
KW   Signal; Repeat; EGF-like domain; Calcium; Immunoglobulin domain;
KW   Hyaluronic acid; Alternative splicing.
FT   SIGNAL        1     20       POTENTIAL.
FT   CHAIN        21   2738       VERSICAN CORE PROTEIN.
FT   NON_CONS    348    349
FT   DOMAIN       21    146       IG-LIKE V-TYPE.
FT   DOMAIN      167    244       LINK 1.
FT   DOMAIN      265    346       LINK 2.
FT   DOMAIN     <349    695       GAG-ALPHA
FT                                (GLUCOSAMINOGLYCAN ATTACHMENT DOMAIN).
FT   DOMAIN      696   2431       GAG-BETA.
FT   DOMAIN     2431   2467       EGF-LIKE 1.
FT   DOMAIN     2469   2505       EGF-LIKE 2, CALCIUM-BINDING.
FT   DOMAIN     2518   2632       C-TYPE LECTIN.
FT   DOMAIN     2637   2695       SUSHI.
FT   DISULFID     44    130       BY SIMILARITY.
FT   DISULFID    172    243       BY SIMILARITY.
FT   DISULFID    196    217       BY SIMILARITY.
FT   DISULFID    270    345       BY SIMILARITY.
FT   DISULFID    294    315       BY SIMILARITY.
FT   DISULFID   2435   2446       BY SIMILARITY.
FT   DISULFID   2440   2455       BY SIMILARITY.
FT   DISULFID   2457   2466       BY SIMILARITY.
FT   DISULFID   2473   2484       BY SIMILARITY.
FT   DISULFID   2478   2493       BY SIMILARITY.
FT   DISULFID   2495   2504       BY SIMILARITY.
FT   DISULFID   2511   2522       BY SIMILARITY.
FT   DISULFID   2539   2631       BY SIMILARITY.
FT   DISULFID   2607   2623       BY SIMILARITY.
FT   DISULFID   2638   2681       BY SIMILARITY.
FT   DISULFID   2667   2694       BY SIMILARITY.
FT   CARBOHYD     57     57       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    330    330       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    692    692       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    758    758       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    802    802       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    805    805       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1257   1257       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1435   1435       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1633   1633       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1660   1660       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1684   1684       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1738   1738       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1848   1848       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   2004   2004       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   2409   2409       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   2711   2711       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   2721   2721       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   VARSPLIC    349   2431       Missing (in isoform V3).
FT                                /FTId=VSP_003091.
FT   VARSPLIC   2697   2738       PSAYQRTYSKKYLKNSSSVKDNSINTSKHEHRWSRRWQETR
FT                                R -> RKWSFRKNGQPCFNKY (in isoform Vint).
FT                                /FTId=VSP_003092.
FT   CONFLICT   2535   2539       AEREC -> NSARG (IN REF. 4).
SQ   SEQUENCE   2738 AA;  300004 MW;  12CA626D58BD8C6A CRC64;

  Query Match             15.6%;  Score 174;  DB 1;  Length 2738;
  Best Local Similarity   28.5%;  Pred. No. 3.7e-07;
  Matches   47;  Conservative   23;  Mismatches   55;  Indels   40;  Gaps    8;

Qy         17 QRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDF- 75
              |  |||  ||   : |  :: |:  ||  |  | || | :||  : :          |:
Db       2519 QGQCYK--YF---AHRRTWDAAERECRLQGAHLTSILSHEEQMFVNRV-------GHDYQ 2566

Qy         76 WIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEP----SCGSEVCVVMYHQPSAP 131
              ||||        |    :  : |||||  |: ||  ::|    | | :  |:::|:
Db       2567 WIGL--------NDKMFEHDFRWTDGSALQYENWRPNQPDSFFSAGEDCVVIIWHENG-- 2616
```

```
Qy        132 AGIGGPYMFQWNDDRCNMKNNFICKYS----DEKPAVPSREAEGE 172
                  ||||  ||    : ||       : | | : :  |:
Db       2617 ---------QWNDVPCNYHLTYTCKKGTVACGQPPVVENAKTFGK 2652


RESULT 9
PGCV_MOUSE
ID   PGCV_MOUSE     STANDARD;      PRT;  3358 AA.
AC   Q62059; Q62058; Q9CUU0;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   Versican core protein precursor (Large fibroblast proteoglycan)
DE   (Chondroitin sulfate proteoglycan core protein 2) (PG-M).
GN   CSPG2.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORMS V0; V1 AND V2).
RC   STRAIN=C57BL/6, and Swiss Webster; TISSUE=Brain;
RX   MEDLINE=95122551; PubMed=7822336;
RA   Ito K., Shinomura T., Zako M., Ujita M., Kimata K.;
RT   "Multiple forms of mouse PG-M, a large chondroitin sulfate
RT   proteoglycan generated by alternative splicing.";
RL   J. Biol. Chem. 270:958-965(1995).
RN   [2]
RP   SEQUENCE FROM N.A. (ISOFORM V3).
RC   STRAIN=C57BL/6;
RX   MEDLINE=95181355; PubMed=7876137;
RA   Zako M., Shinomura T., Ujita M., Ito K., Kimata K.;
RT   "Expression of PG-M(V3), an alternatively spliced form of PG-M
RT   without a chondroitin sulfate attachment in region in mouse and human
RT   tissues.";
RL   J. Biol. Chem. 270:3914-3918(1995).
RN   [3]
RP   SEQUENCE OF 1-1692 FROM N.A. (ISOFORM V1).
RC   STRAIN=C57BL/6J; TISSUE=Skin;
RX   MEDLINE=22354683; PubMed=12466851;
RA   Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
RA   Nikaido I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,
RA   Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,
RA   Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
RA   Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
RA   Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.E., Cousins S.,
RA   Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
RA   Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
RA   Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
RA   Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,
RA   Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
RA   Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
RA   Nagashima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G.,
RA   Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
RA   Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
RA   Sandelin A., Schneider C., Semple C.A., Setou M., Shimada K.,
RA   Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
RA   Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,
RA   Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L.,
RA   Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
RA   Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,
RA   Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
RA   Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
RA   Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
RA   Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
RA   Birney E., Hayashizaki Y.;
RT   "Analysis of the mouse transcriptome based on functional annotation of
RT   60,770 full-length cDNAs.";
RL   Nature 420:563-573(2002).
RN   [4]
RP   INTERACTION WITH FBLN1.
RX   MEDLINE=99329059; PubMed=10400671;
RA   Aspberg A., Adam S., Kostka G., Timpl R., Heinegaard D.;
RT   "Fibulin-1 is a ligand for the C-type lectin domains of aggrecan and
RT   versican.";
RL   J. Biol. Chem. 274:20444-20449(1999).
CC   -!- FUNCTION: May play a role in intercellular signaling and in
CC       connecting cells with the extracellular matrix. May take part in
CC       the regulation of cell motility, growth and differentiation. Binds
CC       hyaluronic acid.
CC   -!- SUBUNIT: Interacts with FBLN1.
CC   -!- SUBCELLULAR LOCATION: Secreted; extracellular matrix.
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=4;
CC         Comment=Additional isoforms seem to exist;
CC       Name=V0;
CC         IsoId=Q62059-1; Sequence=Displayed;
CC       Name=V1;
CC         IsoId=Q62059-2; Sequence=VSP_003087, VSP_003088;
CC       Name=V2;
CC         IsoId=Q62059-3; Sequence=VSP_003089;
CC       Name=V3;
CC         IsoId=Q62059-4; Sequence=VSP_003087, VSP_003090;
CC   -!- TISSUE SPECIFICITY: V2 is found only in brain.
CC   -!- DEVELOPMENTAL STAGE: Disappears after the cartilage development.
CC   -!- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
CC   -!- SIMILARITY: Contains 2 link domains.
CC   -!- SIMILARITY: Contains 2 EGF-like domains.
CC   -!- SIMILARITY: Contains 1 C-type lectin family domain.
CC   -!- SIMILARITY: Contains 1 Sushi (SCR) domain.
CC   -!- SIMILARITY: Belongs to the aggrecan/versican proteoglycan family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; D16263; BAA03796.1; -.
DR   EMBL; D28599; -; NOT_ANNOTATED_CDS.
DR   EMBL; D32040; BAA06802.1; -.
DR   EMBL; AK014525; BAB29411.2; -.
DR   HSSP; P01132; 1EPG.
DR   MGD; MGI:102889; Cspg2.
DR   InterPro; IPR000152; Asx_hydroxyl_S.
DR   InterPro; IPR000742; EGF_2.
DR   InterPro; IPR001881; EGF_Ca.
DR   InterPro; IPR006209; EGF_like.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003599; Ig.
DR   InterPro; IPR001304; Lectin_C.
DR   InterPro; IPR000538; Link.
DR   InterPro; IPR000436; Sushi_SCR_CCP.
DR   Pfam; PF00008; EGF; 2.
DR   Pfam; PF00047; ig; 1.
DR   Pfam; PF00059; lectin_c; 1.
DR   Pfam; PF00084; sushi; 1.
DR   Pfam; PF00193; Xlink; 2.
DR   PRINTS; PR01265; LINKMODULE.
DR   ProDom; PD000918; Link; 2.
DR   SMART; SM00032; CCP; 1.
DR   SMART; SM00034; CLECT; 1.
DR   SMART; SM00179; EGF_CA; 1.
DR   SMART; SM00409; IG; 1.
DR   SMART; SM00445; LINK; 2.
DR   PROSITE; PS00010; ASX_HYDROXYL; 1.
DR   PROSITE; PS00615; C_TYPE_LECTIN_1; 1.
DR   PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
DR   PROSITE; PS00022; EGF_1; 2.
DR   PROSITE; PS01186; EGF_2; 1.
DR   PROSITE; PS50026; EGF_3; 2.
DR   PROSITE; PS01187; EGF_CA; 1.
DR   PROSITE; PS50835; IG_LIKE; 1.
DR   PROSITE; PS01241; LINK; 2.
```

```
KW   Glycoprotein; Proteoglycan; Lectin; Extracellular matrix; Sushi;
KW   Signal; Repeat; EGF-like domain; Calcium; Immunoglobulin domain;
KW   Hyaluronic acid; Alternative splicing.
FT   SIGNAL        1     20       POTENTIAL.
FT   CHAIN        21   3358       VERSICAN CORE PROTEIN.
FT   DOMAIN       21    146       IG-LIKE V-TYPE.
FT   DOMAIN      167    244       LINK 1.
FT   DOMAIN      265    346       LINK 2.
FT   DOMAIN      348   1308       GAG-ALPHA
FT                                (GLUCOSAMINOGLYCAN ATTACHMENT DOMAIN).
FT   DOMAIN     1309   3052       GAG-BETA.
FT   DOMAIN     3052   3088       EGF-LIKE 1.
FT   DOMAIN     3090   3126       EGF-LIKE 2, CALCIUM-BINDING (POTENTIAL).
FT   DOMAIN     3139   3253       C-TYPE LECTIN.
FT   DOMAIN     3258   3316       SUSHI.
FT   DISULFID     44    130       BY SIMILARITY.
FT   DISULFID    172    243       BY SIMILARITY.
FT   DISULFID    196    217       BY SIMILARITY.
FT   DISULFID    270    333       BY SIMILARITY.
FT   DISULFID    294    315       BY SIMILARITY.
FT   DISULFID   3056   3067       BY SIMILARITY.
FT   DISULFID   3061   3076       BY SIMILARITY.
FT   DISULFID   3078   3087       BY SIMILARITY.
FT   DISULFID   3094   3105       BY SIMILARITY.
FT   DISULFID   3099   3114       BY SIMILARITY.
FT   DISULFID   3116   3125       BY SIMILARITY.
FT   DISULFID   3132   3143       BY SIMILARITY.
FT   DISULFID   3160   3252       BY SIMILARITY.
FT   DISULFID   3228   3244       BY SIMILARITY.
FT   DISULFID   3259   3302       BY SIMILARITY.
FT   DISULFID   3288   3315       BY SIMILARITY.
FT   CARBOHYD     57     57       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    330    330       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    351    351       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    441    441       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    807    807       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    914    914       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    951    951       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1305   1305       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1372   1372       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1679   1679       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   2054   2054       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   2244   2244       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   2362   2362       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   2627   2627       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   3030   3030       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   3332   3332       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   3342   3342       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   VARSPLIC    348    348       P -> R (in isoform V1 and isoform V3).
FT                                /FTId=VSP_003087.
FT   VARSPLIC    349   1308       Missing (in isoform V1).
FT                                /FTId=VSP_003088.
FT   VARSPLIC   1309   3052       Missing (in isoform V2).
FT                                /FTId=VSP_003089.
FT   VARSPLIC    349   3052       Missing (in isoform V3).
FT                                /FTId=VSP_003090.
FT   CONFLICT    126    126       A -> G (IN REF. 3).
FT   CONFLICT    348    348       MISSING (IN REF. 3).
FT   CONFLICT   1658   1658       I -> T (IN REF. 3).
FT   CONFLICT   1674   1680       TVWNSNS -> QFGIQTA (IN REF. 3).
SQ   SEQUENCE   3358 AA;  366938 MW;  071B80026BC0762D CRC64;

  Query Match             15.6%;  Score 174;  DB 1;  Length 3358;
  Best Local Similarity   28.5%;  Pred. No. 4.7e-07;
  Matches   47;  Conservative   23;  Mismatches   55;  Indels   40;  Gaps    8;

Qy         17 QRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDF- 75
              |  |||  ||   : |  :: |:  ||  |  | || | :||  : :         |:
Db       3140 QGQCYK--YF---AHRRTWDAAERECRLQGAHLTSILSHEEQMFVNRV-------GHDYQ 3187

Qy         76 WIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEP----SCGSEVCVVMYHQPSAP 131
              ||||        |    :  : |||||  |: ||  ::|    | | :  |:::|:
Db       3188 WIGL--------NDKMFEHDFRWTDGSALQYENWRPNQPDSFFSAGEDCVVIIWHENG-- 3237

Qy        132 AGIGGPYMFQWNDDRCNMKNNFICKYS----DEKPAVPSREAEGE 172
              ||||  ||    : ||       : | | : :  |:
Db       3238 ---------QWNDVPCNYHLTYTCKKGTVACGQPPVVENAKTFGK 3273


RESULT 10
PGCV_HUMAN
ID   PGCV_HUMAN     STANDARD;      PRT;  3396 AA.
AC   P13611; P20754; Q13010; Q13189; Q15123; Q9UNW5;
DT   01-JAN-1990 (Rel. 13, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Versican core protein precursor (Large fibroblast proteoglycan)
DE   (Chondroitin sulfate proteoglycan core protein 2) (PG-M) (Glial
DE   hyaluronate-binding protein) (GHAP).
GN   CSPG2.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORM V0).
RX   MEDLINE=95105188; PubMed=7528742;
RA   Naso M.F., Zimmermann D.R., Iozzo R.V.;
RT   "Characterization of the complete genomic structure of the human
RT   versican gene and functional analysis of its promoter.";
RL   J. Biol. Chem. 269:32999-33008(1994).
RN   [2]
RP   SEQUENCE FROM N.A. (ISOFORM V1).
RC   TISSUE=Placenta;
RX   MEDLINE=90059882; PubMed=2583089;
RA   Zimmermann D.R., Ruoslahti E.;
RT   "Multiple domains of the large fibroblast proteoglycan, versican.";
RL   EMBO J. 8:2975-2981(1989).
RN   [3]
RP   SEQUENCE FROM N.A. (ISOFORM V2).
RC   TISSUE=Glial tumor;
RX   MEDLINE=95105187; PubMed=7806529;
RA   Dours-Zimmermann M.T., Zimmermann D.R.;
RT   "A novel glycosaminoglycan attachment domain identified in two
RT   alternative splice variants of human versican.";
RL   J. Biol. Chem. 269:32992-32998(1994).
RN   [4]
RP   SEQUENCE OF 2711-3396 FROM N.A.
RC   TISSUE=Lung fibroblast;
RX   MEDLINE=88007514; PubMed=2820964;
RA   Krusius T., Gehlsen K.R., Ruoslahti E.;
RT   "A fibroblast chondroitin sulfate proteoglycan core protein contains
RT   lectin-like and growth factor-like sequences.";
RL   J. Biol. Chem. 262:13120-13125(1987).
RN   [5]
RP   SEQUENCE OF 251-347 FROM N.A.
RX   MEDLINE=93122792; PubMed=1478664;
RA   Iozzo R.V., Naso M.F., Cannizzaro L.A., Wasmuth J.J.,
RA   McPherson J.D.;
RT   "Mapping of the versican proteoglycan gene (CSPG2) to the long arm of
RT   human chromosome 5 (5q12-5q14).";
RL   Genomics 14:845-851(1992).
RN   [6]
RP   SEQUENCE FROM N.A. (ISOFORM V3).
RC   TISSUE=Brain;
RX   MEDLINE=95181355; PubMed=7876137;
RA   Zako M., Shinomura T., Ujita M., Ito K., Kimata K.;
RT   "Expression of PG-M(V3), an alternatively spliced form of PG-M
RT   without a chondroitin sulfate attachment in region in mouse and human
RT   tissues.";
RL   J. Biol. Chem. 270:3914-3918(1995).
RN   [7]
RP   SEQUENCE OF 3333-3396 FROM N.A. (ISOFORM VINT).
RC   TISSUE=Aortic smooth muscle;
RX   MEDLINE=99327053; PubMed=10397680;
RA   Lemire J.M., Braun K.R., Maurel P., Kaplan E.D., Schwartz S.M.,
```

```
RA   Wight T.N.;
RT   "Versican/PG-M isoforms in vascular smooth muscle cells.";
RL   Arterioscler. Thromb. Vasc. Biol. 19:1630-1639(1999).
RN   [8]
RP   PARTIAL SEQUENCE.
RC   TISSUE=Brain;
RX   MEDLINE=89174663; PubMed=2466833;
RA   Perides G., Lane W.S., Andrews D., Dahl D., Bignami A.;
RT   "Isolation and partial characterization of a glial
RT   hyaluronate-binding protein.";
RL   J. Biol. Chem. 264:5981-5987(1989).
RN   [9]
RP   TISSUE SPECIFICITY OF ISOFORMS.
RX   MEDLINE=96213482; PubMed=8627343;
RA   Paulus W., Baur I., Dours-Zimmermann M.T., Zimmermann D.R.;
RT   "Differential expression of versican isoforms in brain tumors.";
RL   J. Neuropathol. Exp. Neurol. 55:528-533(1996).
CC   -!- FUNCTION: May play a role in intercellular signaling and in
CC       connecting cells with the extracellular matrix. May take part in
CC       the regulation of cell motility, growth and differentiation. Binds
CC       hyaluronic acid.
CC   -!- SUBUNIT: Interacts with FBLN1 (By similarity).
CC   -!- SUBCELLULAR LOCATION: Secreted; extracellular matrix.
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=5;
CC         Comment=Additional isoforms seem to exist;
CC       Name=V0;
CC         IsoId=P13611-1; Sequence=Displayed;
CC       Name=V1;
CC         IsoId=P13611-2; Sequence=VSP_003082, VSP_003083;
CC       Name=V2;
CC         IsoId=P13611-3; Sequence=VSP_003084;
CC       Name=V3;
CC         IsoId=P13611-4; Sequence=VSP_003082, VSP_003085;
CC       Name=Vint;
CC         IsoId=P13611-5; Sequence=VSP_003086;
CC   -!- TISSUE SPECIFICITY: Cerebral white matter. V0 and V1 is expressed
CC       in normal brain, gliomas, medulloblastomas, schwannomas,
CC       neurofibromas, and menningiomas; v2 is restricted to normal brain
CC       and gliomas; v3 is found in all these tissues except
CC       medulloblastomas.
CC   -!- DEVELOPMENTAL STAGE: Disappears after the cartilage development.
CC   -!- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
CC   -!- SIMILARITY: Contains 2 link domains.
CC   -!- SIMILARITY: Contains 2 EGF-like domains.
CC   -!- SIMILARITY: Contains 1 C-type lectin family domain.
CC   -!- SIMILARITY: Contains 1 Sushi (SCR) domain.
CC   -!- SIMILARITY: Belongs to the aggrecan/versican proteoglycan family.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; U16306; AAA65018.1; -.
DR   EMBL; X15998; CAA34128.1; -.
DR   EMBL; S52488; AAB24878.1; -.
DR   EMBL; U26555; AAA67565.1; -.
DR   EMBL; D32039; BAA06801.1; -.
DR   EMBL; J02814; AAA36437.1; -.
DR   EMBL; AF084545; AAD48545.1; -.
DR   PIR; S06014; A60979.
DR   HSSP; P01132; 1EGF.
DR   Genew; HGNC:2464; CSPG2.
DR   MIM; 118661; -.
DR   GO; GO:0005578; C:extracellular matrix; TAS.
DR   GO; GO:0005540; F:hyaluronic acid binding; TAS.
DR   GO; GO:0008037; P:cell recognition; TAS.
DR   GO; GO:0007275; P:development; TAS.
DR   InterPro; IPR000152; Asx_hydroxyl_S.
DR   InterPro; IPR000742; EGF_2.
DR   InterPro; IPR001881; EGF_Ca.
DR   InterPro; IPR006209; EGF_like.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003599; Ig.
DR   InterPro; IPR001304; Lectin_C.
DR   InterPro; IPR000538; Link.
DR   InterPro; IPR000436; Sushi_SCR_CCP.
DR   Pfam; PF00008; EGF; 2.
DR   Pfam; PF00059; lectin_c; 1.
DR   Pfam; PF00084; sushi; 1.
DR   Pfam; PF00193; Xlink; 2.
DR   PRINTS; PR01265; LINKMODULE.
DR   ProDom; PD000918; Link; 2.
DR   SMART; SM00032; CCP; 1.
DR   SMART; SM00034; CLECT; 1.
DR   SMART; SM00179; EGF_CA; 1.
DR   SMART; SM00409; IG; 1.
DR   SMART; SM00445; LINK; 2.
DR   PROSITE; PS00010; ASX_HYDROXYL; 1.
DR   PROSITE; PS00615; C_TYPE_LECTIN_1; 1.
DR   PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
DR   PROSITE; PS00022; EGF_1; 2.
DR   PROSITE; PS01186; EGF_2; 1.
DR   PROSITE; PS50026; EGF_3; 2.
DR   PROSITE; PS01187; EGF_CA; 1.
DR   PROSITE; PS50835; IG_LIKE; 1.
DR   PROSITE; PS01241; LINK; 2.
KW   Glycoprotein; Proteoglycan; Lectin; Extracellular matrix; Sushi;
KW   Signal; Repeat; EGF-like domain; Calcium; Immunoglobulin domain;
KW   Hyaluronic acid; Alternative splicing.
FT   SIGNAL        1     20       POTENTIAL.
FT   CHAIN        21   3396       VERSICAN CORE PROTEIN.
FT   DOMAIN       21    146       IG-LIKE V-TYPE.
FT   DOMAIN      167    244       LINK 1.
FT   DOMAIN      265    346       LINK 2.
FT   DOMAIN      348   1335       GAG-ALPHA
FT                                (GLUCOSAMINOGLYCAN ATTACHMENT DOMAIN).
FT   DOMAIN     1336   3089       GAG-BETA.
FT   DOMAIN     3089   3125       EGF-LIKE 1.
FT   DOMAIN     3127   3163       EGF-LIKE 2, CALCIUM-BINDING (POTENTIAL).
FT   DOMAIN     3176   3290       C-TYPE LECTIN.
FT   DOMAIN     3295   3353       SUSHI.
FT   DISULFID     44    130       BY SIMILARITY.
FT   DISULFID    172    243       BY SIMILARITY.
FT   DISULFID    196    217       BY SIMILARITY.
FT   DISULFID    270    345       BY SIMILARITY.
FT   DISULFID    294    315       BY SIMILARITY.
FT   DISULFID   3093   3104       BY SIMILARITY.
FT   DISULFID   3098   3113       BY SIMILARITY.
FT   DISULFID   3115   3124       BY SIMILARITY.
FT   DISULFID   3131   3142       BY SIMILARITY.
FT   DISULFID   3136   3151       BY SIMILARITY.
FT   DISULFID   3153   3162       BY SIMILARITY.
FT   DISULFID   3169   3180       BY SIMILARITY.
FT   DISULFID   3197   3289       BY SIMILARITY.
FT   DISULFID   3265   3281       BY SIMILARITY.
FT   DISULFID   3296   3339       BY SIMILARITY.
FT   DISULFID   3325   3352       BY SIMILARITY.
FT   CARBOHYD     57     57       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    330    330       N-LINKED (GLCNAC. . .).
FT   CARBOHYD    615    615       N-LINKED (GLCNAC. . .) (POTENTIAL).

  Query Match             15.6%;  Score 174;  DB 1;  Length 3396;
  Best Local Similarity   28.5%;  Pred. No. 4.8e-07;
  Matches   47;  Conservative   23;  Mismatches   55;  Indels   40;  Gaps    8;

Qy         17 QRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDF- 75
              |  |||  ||   :  |  :: |:  ||  |  |  || | :||  : :          |:
Db       3177 QGQCYK--YF---AHRRTWDAAERECRLQGAHLTSILSHEEQMFVNRV-------GHDYQ 3224

Qy         76 WIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEP----SCGSEVCVVMYHQPSAP 131
                ||||        |    :  : |||||  |: ||  ::|    | | :  |:::|:
```

```
Db       3225 WIGL--------NDKMFEHDFRWTDGSTLQYENWRPNQPDSFFSAGEDCVVIIWHENG-- 3274

Qy        132 AGIGGPYMFQWNDDRCNMKNNFICKYS----DEKPAVPSREAEGE 172
                  ||||  ||    : ||       : | | : :  |:
Db       3275 ---------QWNDVPCNYHLTYTCKKGTVACGQPPVVENAKTFGK 3310


RESULT 11
PGCV_CHICK
ID   PGCV_CHICK     STANDARD;      PRT;  3562 AA.
AC   Q90953; Q90945;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Versican core protein precursor (Large fibroblast proteoglycan)
DE   (Chondroitin sulfate proteoglycan core protein 2) (PG-M).
GN   CPSG2.
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
OX   NCBI_TaxID=9031;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORMS V0 AND V1).
RC   STRAIN=White leghorn; TISSUE=Limb bud;
RX   MEDLINE=93300846; PubMed=8314802;
RA   Shinomura T., Nishida Y., Ito K., Kimata K.;
RT   "cDNA cloning of PG-M, a large chondroitin sulfate proteoglycan
RT   expressed during chondrogenesis in chick limb buds. Alternative
RT   spliced multiforms of PG-M and their relationships to versican.";
RL   J. Biol. Chem. 268:14461-14469(1993).
CC   -!- FUNCTION: May play a role in intercellular signaling and in
CC       connecting cells with the extracellular matrix. May take part in
CC       the regulation of cell motility, growth and differentiation. Binds
CC       hyaluronic acid.
CC   -!- SUBCELLULAR LOCATION: Secreted; extracellular matrix.
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=2;
CC         Comment=Additional isoforms seem to exist;
CC       Name=V0;
CC         IsoId=Q90953-1; Sequence=Displayed;
CC       Name=V1;
CC         IsoId=Q90953-2; Sequence=VSP_003093;
CC   -!- TISSUE SPECIFICITY: Prechondrogenic condensation area of
CC       developing limb buds.
CC   -!- DEVELOPMENTAL STAGE: Disappears after the cartilage development
CC       (By similarity).
CC   -!- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
CC   -!- SIMILARITY: Contains 2 link domains.
CC   -!- SIMILARITY: Contains 2 EGF-like domains.
CC   -!- SIMILARITY: Contains 1 C-type lectin family domain.
CC   -!- SIMILARITY: Contains 1 Sushi (SCR) domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X60226; CAA42787.1; -.
DR   EMBL; D13542; BAA02742.1; -.
DR   PIR; A47171; A47171.
DR   HSSP; P00740; 1EDM.
DR   InterPro; IPR000152; Asx_hydroxyl_S.
DR   InterPro; IPR000742; EGF_2.
DR   InterPro; IPR001881; EGF_Ca.
DR   InterPro; IPR006209; EGF_like.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003599; Ig.
DR   InterPro; IPR001304; Lectin_C.
DR   InterPro; IPR000538; Link.
DR   InterPro; IPR000436; Sushi_SCR_CCP.
DR   Pfam; PF00008; EGF; 2.
DR   Pfam; PF00047; ig; 1.
DR   Pfam; PF00059; lectin_c; 1.
DR   Pfam; PF00084; sushi; 1.
DR   Pfam; PF00193; Xlink; 2.
DR   PRINTS; PR01265; LINKMODULE.
DR   ProDom; PD000918; Link; 2.
DR   SMART; SM00032; CCP; 1.
DR   SMART; SM00034; CLECT; 1.
DR   SMART; SM00179; EGF_CA; 1.
DR   SMART; SM00409; IG; 1.
DR   SMART; SM00445; LINK; 2.
DR   PROSITE; PS00010; ASX_HYDROXYL; 1.
DR   PROSITE; PS00615; C_TYPE_LECTIN_1; 1.
DR   PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
DR   PROSITE; PS00022; EGF_1; 2.
DR   PROSITE; PS01186; EGF_2; 1.
DR   PROSITE; PS50026; EGF_3; 2.
DR   PROSITE; PS01187; EGF_CA; 1.
DR   PROSITE; PS50835; IG_LIKE; 1.
DR   PROSITE; PS01241; LINK; 2.
KW   Glycoprotein; Proteoglycan; Lectin; Extracellular matrix; Sushi;
KW   Signal; Repeat; EGF-like domain; Calcium; Immunoglobulin domain;
KW   Hyaluronic acid; Alternative splicing.
FT   SIGNAL        1     26       POTENTIAL.
FT   CHAIN        27   3562       VERSICAN CORE PROTEIN.
FT   DOMAIN       27    143       IG-LIKE V-TYPE.
FT   DOMAIN      166    243       LINK 1.
FT   DOMAIN      264    345       LINK 2.
FT   DOMAIN     3254   3290       EGF-LIKE 1.
FT   DOMAIN     3292   3328       EGF-LIKE 2, CALCIUM-BINDING (POTENTIAL).
FT   DOMAIN     3341   3455       C-TYPE LECTIN.
FT   DOMAIN     3460   3518       SUSHI.
FT   DISULFID     44    129       BY SIMILARITY.
FT   DISULFID    171    242       BY SIMILARITY.
FT   DISULFID    195    216       BY SIMILARITY.
FT   DISULFID    269    344       BY SIMILARITY.
FT   DISULFID    293    314       BY SIMILARITY.
FT   DISULFID   3258   3269       BY SIMILARITY.
FT   DISULFID   3263   3278       BY SIMILARITY.
FT   DISULFID   3280   3289       BY SIMILARITY.
FT   DISULFID   3296   3307       BY SIMILARITY.
FT   DISULFID   3301   3316       BY SIMILARITY.
FT   DISULFID   3318   3327       BY SIMILARITY.
FT   DISULFID   3334   3345       BY SIMILARITY.
FT   DISULFID   3362   3454       BY SIMILARITY.
FT   DISULFID   3430   3446       BY SIMILARITY.
FT   DISULFID   3461   3504       BY SIMILARITY.
FT   DISULFID   3490   3517       BY SIMILARITY.
FT   CARBOHYD    163    163       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    235    235       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    329    329       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    529    529       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    709    709       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    948    948       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1409   1409       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1479   1479       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1523   1523       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1530   1530       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1625   1625       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1751   1751       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1988   1988       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   2088   2088       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   2089   2089       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   2507   2507       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   2642   2642       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   2679   2679       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   2748   2748       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   2762   2762       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   3069   3069       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   3194   3194       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   3232   3232       N-LINKED (GLCNAC. . .) (POTENTIAL).
```

```
FT   CARBOHYD   3545   3545       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   VARSPLIC    485   1411       Missing (in isoform V1).
FT                                /FTId=VSP_003093.
SQ   SEQUENCE   3562 AA;  388078 MW;  9BC566E88C1602D2 CRC64;

  Query Match             15.3%;  Score 171;  DB 1;  Length 3562;
  Best Local Similarity   28.5%;  Pred. No. 9.5e-07;
  Matches   47;  Conservative   23;  Mismatches   55;  Indels   40;  Gaps    8;

Qy         17 QRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDF- 75
              |  |||  ||   : |  :: |:  ||  |  | || | :||  : :          |:
Db       3342 QGQCYK--YF---AHRRTWDTAERECRLQGAHLTSILSHEEQVFVNRI-------GHDYQ 3389

Qy         76 WIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEP----SCGSEVCVVMYHQPSAP 131
              ||||         |    :  : |||||  |: ||  ::|    | | :  |:::|:
Db       3390 WIGL--------NDKMFERDFRWTDGSPLQYENWRPNQPDSFFSAGEDCVVIIWHENG-- 3439

Qy        132 AGIGGPYMFQWNDDRCNMKNNFICKYS----DEKPAVPSREAEGE 172
                ||||  ||    : ||       : | | : :  |:
Db       3440 ---------QWNDVPCNYHLTYTCKKGTVACGQPPVVENAKTFGK 3475


RESULT 12
CD93_RAT
ID   CD93_RAT       STANDARD;      PRT;   643 AA.
AC   Q9ET61; Q9JIZ6;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Complement component C1q receptor precursor (Complement component 1, q
DE   subcomponent, receptor 1) (C1qRp) (C1qR(p)) (C1q/MBL/SPA receptor)
DE   (CD93 antigen) (Cell surface antigen AA4).
GN   C1QR1 OR CD93 OR C1QRP.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=PVG; TISSUE=Natural killer cells;
RX   MEDLINE=20545218; PubMed=11093152;
RA   Lovik G., Vaage J.T., Dissen E., Szpirer C., Ryan J.C., Rolstad B.;
RT   "Characterization and molecular cloning of rat C1qRp, a receptor on NK
RT   cells.";
RL   Eur. J. Immunol. 30:3355-3362(2000).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Wistar; TISSUE=Lung;
RX   MEDLINE=20507883; PubMed=10934210;
RA   Dean Y.D., McGreal E.P., Akatsu H., Gasque P.;
RT   "Molecular and cellular properties of the rat AA4 antigen, a C-type
RT   lectin-like receptor with structural homology to thrombomodulin.";
RL   J. Biol. Chem. 275:34382-34392(2000).
CC   -!- FUNCTION: Receptor (or element of a larger receptor complex) for
CC       C1q, mannose-binding lectin (MBL2) and pulmonary surfactant
CC       protein A (SPA). May mediate the enhancement of phagocytosis in
CC       monocytes and macrophages upon interaction with soluble defense
CC       collagens. May play a role in intercellular adhesion.
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC   -!- TISSUE SPECIFICITY: Widely expressed. Highly expressed in lung and
CC       heart. Expressed at lower level in brain, thymus, liver, spleen,
CC       intestine, kidney, adrenal gland, muscle and testis. Expressed on
CC       endothelial cells, platelets, undifferentiated monocytes and
CC       circulating natural killer cells.
CC   -!- PTM: N- and O-glycosylated (By similarity).
CC   -!- SIMILARITY: Contains 1 C-type lectin family domain.
CC   -!- SIMILARITY: Contains 5 EGF-like domains.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; AF136537; AAG01572.1; -.
DR   EMBL; AF160978; AAF80402.1; -.
DR   HSSP; P35555; 1EMN.
DR   GO; GO:0016021; C:integral to membrane; ISS.
DR   GO; GO:0004872; F:receptor activity; ISS.
DR   GO; GO:0016337; P:cell-cell adhesion; ISS.
DR   GO; GO:0042116; P:macrophage activation; ISS.
DR   GO; GO:0006909; P:phagocytosis; ISS.
DR   InterPro; IPR000152; Asx_hydroxyl_S.
DR   InterPro; IPR001881; EGF_Ca.
DR   InterPro; IPR006209; EGF_like.
DR   InterPro; IPR001304; Lectin_C.
DR   Pfam; PF00008; EGF; 4.
DR   Pfam; PF00059; lectin_c; 1.
DR   SMART; SM00034; CLECT; 1.
DR   SMART; SM00179; EGF_CA; 3.
DR   PROSITE; PS00010; ASX_HYDROXYL; 3.
DR   PROSITE; PS00615; C_TYPE_LECTIN_1; FALSE_NEG.
DR   PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
DR   PROSITE; PS01186; EGF_2; 3.
DR   PROSITE; PS50026; EGF_3; 4.
DR   PROSITE; PS01187; EGF_CA; 3.
KW   Cell adhesion; Receptor; Repeat; Signal; Transmembrane;
KW   EGF-like domain; Lectin; Glycoprotein.
FT   SIGNAL        1     23       POTENTIAL.
FT   CHAIN        24    643       COMPLEMENT COMPONENT C1Q RECEPTOR.
FT   DOMAIN       24    571       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    572    592       POTENTIAL.
FT   DOMAIN      593    643       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN       31    173       C-TYPE LECTIN.
FT   DOMAIN      257    298       EGF-LIKE 1.
FT   DOMAIN      299    341       EGF-LIKE 2.
FT   DOMAIN      342    381       EGF-LIKE 3, CALCIUM-BINDING (POTENTIAL).
FT   DOMAIN      382    423       EGF-LIKE 4, CALCIUM-BINDING (POTENTIAL).
FT   DOMAIN      424    462       EGF-LIKE 5, CALCIUM-BINDING (POTENTIAL).
FT   DISULFID    261    272       BY SIMILARITY.
FT   DISULFID    268    282       BY SIMILARITY.
FT   DISULFID    284    297       BY SIMILARITY.
FT   DISULFID    303    314       BY SIMILARITY.
FT   DISULFID    308    325       BY SIMILARITY.
FT   DISULFID    327    340       BY SIMILARITY.
FT   DISULFID    346    355       BY SIMILARITY.
FT   DISULFID    351    364       BY SIMILARITY.
FT   DISULFID    366    380       BY SIMILARITY.
FT   DISULFID    386    397       BY SIMILARITY.
FT   DISULFID    393    406       BY SIMILARITY.
FT   DISULFID    408    422       BY SIMILARITY.
FT   DISULFID    428    437       BY SIMILARITY.
FT   DISULFID    433    446       BY SIMILARITY.
FT   DISULFID    448    461       BY SIMILARITY.
FT   CARBOHYD    322    322       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    498    498       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CONFLICT    417    417       E -> K (IN REF. 2).
SQ   SEQUENCE   643 AA;  68781 MW;  9AE4C933AD943DB6 CRC64;

  Query Match             14.8%;  Score 165;  DB 1;  Length 643;
  Best Local Similarity   25.2%;  Pred. No. 3.9e-07;
  Matches   52;  Conservative   35;  Mismatches   73;  Indels   46;  Gaps   10;

Qy          3 GRLLSGQP------VCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESED 56
              |:| :|         || |      ||   :      :|:  ||:  |  :|| | :::||:
Db         15 GQLWAGAAADSEAVVCEG---TACYTAHW-----GKLSAAEAQHRCNENGGNLATVKSEE 66

Qy         57 EQKLIEKFIENLL----PSD---GDFWIGLRRREEKQSNSTACQDLYAWT-DGSISQFRN 108
                | : ::: :  ||     ||:   | |||||:| : | :         ::|   |  : : |
Db         67 EARHVQEALAQLLKTKAPSETKIGKFWIGLQREKGKCTYHDLPMKGFSWVGGGEDTTYSN 126

Qy        109 WY-VDEPSCGSEVCVVMY-------HQPSAPAGIGGPYMFQWNDDRC--------NMKNNF 153
              ||   : || |: || :        |    |         :|::  |        |    |
Db        127 WYKASKSSCISKRCVSLILDLSLKPHPSHLP---------KWHESPCGTPDAPGNSIEGF 177
```

```
Qy        154 ICKYSDEKPAVPSREAEGEETELTTP 179
              :||:: :    |       :   |||
Db        178 LCKFNFKGMCSPLALGGPGQLTYTTP 203


RESULT 13
PGCA_BOVIN
ID   PGCA_BOVIN     STANDARD;      PRT;  2364 AA.
AC   P13608; P79117; Q28159;
DT   01-JAN-1990 (Rel. 13, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   Aggrecan core protein precursor (Cartilage-specific proteoglycan core
DE   protein) (CSPCP).
GN   AGC1.
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Hering T.M., Kollar J., Huynh T.D.;
RL   Submitted (SEP-1996) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE OF 563-1056 FROM N.A.
RX   MEDLINE=89380219; PubMed=2528543;
RA   Antonsson P., Heinegaard D., Oldberg A.;
RT   "The keratan sulfate-enriched region of bovine cartilage proteoglycan
RT   consists of a consecutively repeated hexapeptide motif.";
RL   J. Biol. Chem. 264:16170-16173(1989).
RN   [3]
RP   SEQUENCE OF 1609-2113 AND 2151-2364 FROM N.A.
RX   MEDLINE=87270630; PubMed=3111460;
RA   Oldberg A., Antonsson P., Heinegaard D.;
RT   "The partial amino acid sequence of bovine cartilage proteoglycan,
RT   deduced from a cDNA clone, contains numerous Ser-Gly sequences
RT   arranged in homologous repeats.";
RL   Biochem. J. 243:255-259(1987).
RN   [4]
RP   SEQUENCE OF 2114-2150 FROM N.A.
RC   TISSUE=Cartilage;
RX   MEDLINE=93352525; PubMed=8349621;
RA   Fueloep C., Walcz E., Valyon M., Glant T.T.;
RT   "Expression of alternatively spliced epidermal growth factor-like
RT   domains in aggrecans of different species. Evidence for a novel
RT   module.";
RL   J. Biol. Chem. 268:17377-17383(1993).
RN   [5]
RP   PARTIAL SEQUENCE.
RX   MEDLINE=85027710; PubMed=6489519;
RA   Perin J.-P., Bonnet F., Jolles J., Jolles P.;
RT   "Sequence data concerning the protein core of the cartilage
RT   proteoglycan monomers. Characterization of a sequence allowing the
RT   synthesis of an oligonucleotide probe.";
RL   FEBS Lett. 176:37-42(1984).
RN   [6]
RP   PARTIAL SEQUENCE.
RX   MEDLINE=87005253; PubMed=3530809;
RA   Perin J.P., Bonnet F., Jolles P.;
RT   "Structural relationship between link proteins and proteoglycan
RT   monomers.";
RL   FEBS Lett. 206:73-77(1986).
CC   -!- FUNCTION: This proteoglycan is a major component of extracellular
CC       matrix of cartilagenous tissues. A major function of this protein
CC       is to resist compression in cartilage. It binds avidly to
CC       hyaluronic acid via an amino-terminal globular region. May play a
CC       regulatory role in the matrix assembly of the cartilage.
CC   -!- SUBCELLULAR LOCATION: Secreted; extracellular matrix (By
CC       similarity).
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=2;
CC       Name=1;
CC         IsoId=P13608-1; Sequence=Displayed;
CC       Name=2;
CC         IsoId=P13608-2; Sequence=VSP_003072;
CC   -!- DOMAIN: Two globular domains, G1 and G2, comprise the amino
CC       terminus of the proteoglycan, while another globular region, G3,
CC       makes up the C-terminus. G1 contains link domains and thus
CC       consists of three disulfide-bonded loop structures designated as
CC       the A, B, B' motifs. G2 is similar to G1. The keratan sulfate (KS)
CC       and the chondroitin sulfate (CS) attachment domains lie between G2
CC       and G3.
CC   -!- PTM: CONTAINS MOSTLY CHONDROITIN SULFATE, BUT ALSO N-LINKED AND O-
CC       LINKED (ABOUT 40) OLIGOSACCHARIDES.
CC   -!- PTM: THE KERATAN SULFATE CONTENTS DIFFER CONSIDERABLY BETWEEN
CC       ADULT AND FETAL BOVINE PROTEOGLYCANS.
CC   -!- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
CC   -!- SIMILARITY: Contains 4 link domains.
CC   -!- SIMILARITY: Contains 1 EGF-like domain.
CC   -!- SIMILARITY: Contains 1 C-type lectin family domain.
CC   -!- SIMILARITY: Contains 1 Sushi (SCR) domain.
CC   -!- SIMILARITY: Belongs to the aggrecan/versican proteoglycan family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U76615; AAB38524.1; -.
DR   EMBL; L07053; -; NOT_ANNOTATED_CDS.
DR   PIR; A34234; A39808.
DR   PIR; T42630; T42630.
DR   HSSP; P08709; 1BF9.
DR   InterPro; IPR002353; AntifreezeII.
DR   InterPro; IPR000152; Asx_hydroxyl_S.
DR   InterPro; IPR000742; EGF_2.
DR   InterPro; IPR001881; EGF_Ca.
DR   InterPro; IPR006209; EGF_like.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR001304; Lectin_C.
DR   InterPro; IPR000538; Link.
DR   InterPro; IPR003324; SGXXSG.
DR   InterPro; IPR000436; Sushi_SCR_CCP.
DR   Pfam; PF00008; EGF; 1.
DR   Pfam; PF00047; ig; 1.
DR   Pfam; PF00059; lectin_c; 1.
DR   Pfam; PF02339; SGXXSG; 61.
DR   Pfam; PF00084; sushi; 1.
DR   Pfam; PF00193; Xlink; 4.
DR   PRINTS; PR00356; ANTIFREEZEII.
DR   PRINTS; PR01265; LINKMODULE.
DR   ProDom; PD000918; Link; 4.
DR   SMART; SM00032; CCP; 1.
DR   SMART; SM00034; CLECT; 1.
DR   SMART; SM00179; EGF_CA; 1.
DR   SMART; SM00445; LINK; 4.
DR   PROSITE; PS00010; ASX_HYDROXYL; 1.
DR   PROSITE; PS00615; C_TYPE_LECTIN_1; 1.
DR   PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
DR   PROSITE; PS00022; EGF_1; 1.
DR   PROSITE; PS50026; EGF_3; 1.
DR   PROSITE; PS01187; EGF_CA; 1.
DR   PROSITE; PS50835; IG_LIKE; 1.
DR   PROSITE; PS00290; IG_MHC; FALSE_NEG.
DR   PROSITE; PS01241; LINK; 4.
KW   Glycoprotein; Proteoglycan; Lectin; Signal; Sushi; EGF-like domain;
KW   Calcium; Alternative splicing; Repeat; Immunoglobulin domain.
FT   SIGNAL       1     16       POTENTIAL.
FT   CHAIN       17   2364       AGGRECAN CORE PROTEIN.
FT   DOMAIN      25    147       IG-LIKE V-TYPE.
```

```
FT   DOMAIN       170    247       LINK 1.
FT   DOMAIN       268    349       LINK 2.
FT   DOMAIN       504    581       LINK 3.
FT   DOMAIN       602    683       LINK 4.
FT   DOMAIN       774    907       23 X 6 AA APPROXIMATE TANDEM REPEATS OF
FT                                 E-[EK]-P-F-P-S.
FT   DOMAIN      1433   2112       CS-2.
FT   DOMAIN      2113   2149       EGF-LIKE, CALCIUM-BINDING (POTENTIAL).
FT   DOMAIN      2114   2364       G3.
FT   DOMAIN      2161   2276       C-TYPE LECTIN.
FT   DOMAIN      2280   2338       SUSHI.
FT   DISULFID      51    133       BY SIMILARITY.
FT   DISULFID     175    246       BY SIMILARITY.
FT   DISULFID     199    220       BY SIMILARITY.
FT   DISULFID     273    348       BY SIMILARITY.
FT   DISULFID     297    318       BY SIMILARITY.
FT   DISULFID     509    580       BY SIMILARITY.
FT   DISULFID     533    554       BY SIMILARITY.
FT   DISULFID     607    682       BY SIMILARITY.
FT   DISULFID     631    652       BY SIMILARITY.
FT   DISULFID    2117   2128       BY SIMILARITY.
FT   DISULFID    2182   2274       BY SIMILARITY.
FT   DISULFID    2250   2266       BY SIMILARITY.
FT   DISULFID    2281   2324       BY SIMILARITY.
FT   DISULFID    2310   2337       BY SIMILARITY.
FT   CARBOHYD     126    126       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     239    239       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     333    333       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     387    387       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     611    611       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     667    667       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   VARSPLIC    2114   2150       Missing (in isoform 2).
FT                                 /FTId=VSP_003072.
SQ   SEQUENCE   2364 AA;  246359 MW;  6FF83763420C3D4C CRC64;

  Query Match             14.2%;  Score 158.5;  DB 1;  Length 2364;
  Best Local Similarity   26.8%;  Pred. No. 7.6e-06;
  Matches   48;  Conservative   29;  Mismatches   59;  Indels   43;  Gaps   10;

Qy          9 QPVCRGG---TQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFI 65
              | :|  |    |  ||:  :| |   |  : :|:  ||:    | || : :||   :|:
Db       2151 QKLCEEGWTKFQGHCYR--HFPD---RATWVDAESQCRKQQSHLSSIVTPEEQ----EFV 2201

Qy         66 ENLLPSDGDF-WIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEP----SCGSEV 120
               |   :  |: ||||        |     :  : |:||   || ||  ::|    : | :
Db       2202 NN---NAQDYQWIGL--------NDKTIEGDFRWSDGHSLQFENWRPNQPDNFFATGEDC 2250

Qy        121 CVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYS----DEKPAVPSREAEGEETE 175
              |:::|:            :|||  || :  | ||       | | |       |:: :
Db       2251 VVMIWHEKG----------EWNDVPCNYQLPFTCKKGTVACGEPPVVEHARIFGQKKD 2298


RESULT 14
PGCA_HUMAN
ID   PGCA_HUMAN     STANDARD;      PRT;  2415 AA.
AC   P16112; Q13650;
DT   01-APR-1990 (Rel. 14, Created)
DT   01-AUG-1992 (Rel. 23, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   Aggrecan core protein precursor (Cartilage-specific proteoglycan core
DE   protein) (CSPCP) (Chondroitin sulfate proteoglycan core protein 1).
GN   AGC1 OR CSPG1.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORM 3).
RC   TISSUE=Chondrocytes;
RX   MEDLINE=91093289; PubMed=1985970;
RA   Doege K.J., Sasaki M., Kimura T., Yamada Y.;
RT   "Complete coding sequence and deduced primary structure of the human
RT   cartilage large aggregating proteoglycan, aggrecan. Human-specific
RT   repeats, and additional alternatively spliced forms.";
RL   J. Biol. Chem. 266:894-902(1991).
RN   [2]
RP   SEQUENCE OF 1778-2415 FROM N.A. (ISOFORM 2).
RC   TISSUE=Chondrocytes;
RA   Dudhia J., Hardingham T.E.;
RL   Submitted (JAN-1990) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE OF 1936-2415 FROM N.A. (ISOFORM 1).
RX   MEDLINE=89380154; PubMed=2789216;
RA   Baldwin C.T., Reginato A.M., Prockop D.J.;
RT   "A new epidermal growth factor-like domain in the human core protein
RT   for the large cartilage-specific proteoglycan. Evidence for
RT   alternative splicing of the domain.";
RL   J. Biol. Chem. 264:15747-15750(1989).
RN   [4]
RP   SEQUENCE OF 764-864 FROM N.A.
RC   TISSUE=Blood;
RX   MEDLINE=95128522; PubMed=7827755;
RA   Barry F.P., Neame P.J., Sasse J., Pearson D.;
RT   "Length variation in the keratan sulfate domain of mammalian
RT   aggrecan.";
RL   Matrix Biol. 14:323-328(1994).
CC   -!- FUNCTION: THIS PROTEOGLYCAN IS A MAJOR COMPONENT OF EXTRACELLULAR
CC       MATRIX OF CARTILAGENOUS TISSUES. A MAJOR FUNCTION OF THIS PROTEIN
CC       IS TO RESIST COMPRESSION IN CARTILAGE. IT BINDS AVIDLY TO
CC       HYALURONIC ACID VIA AN AMINO-TERMINAL GLOBULAR REGION.
CC   -!- SUBUNIT: Interacts with FBLN1 (By similarity).
CC   -!- SUBCELLULAR LOCATION: Secreted; extracellular matrix (By
CC       similarity).
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=3;
CC         Comment=Additional isoforms seem to exist;
CC       Name=1;
CC         IsoId=P16112-1; Sequence=Displayed;
CC       Name=2;
CC         IsoId=P16112-2; Sequence=VSP_003074;
CC       Name=3;
CC         IsoId=P16112-3; Sequence=VSP_003074, VSP_003075;
CC   -!- DOMAIN: Two globular domains, G1 and G2, comprise the amino
CC       terminus of the proteoglycan, while another globular region, G3,
CC       makes up the COOH terminus. G1 contains link domains and thus
CC       consists of three disulfide-bonded loop structures designated as
CC       the A, B, B' motifs. G2 is similar to G1. The keratan sulfate (KS)
CC       and the chondroitin sulfate (CS) attachment domains lie between G2
CC       and G3.
CC   -!- PTM: Contains mostly chondroitin sulfate, but also keratan sulfate
CC       chains, N-linked and O-linked oligosaccharides.
CC   -!- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
CC   -!- SIMILARITY: Contains 4 link domains.
CC   -!- SIMILARITY: Contains 1 EGF-like domain.
CC   -!- SIMILARITY: Contains 1 C-type lectin family domain.
CC   -!- SIMILARITY: Contains 1 Sushi (SCR) domain.
CC   -!- SIMILARITY: Belongs to the aggrecan/versican proteoglycan family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M55172; AAA62824.1; -.
DR   EMBL; J05062; AAA35726.1; -.
DR   EMBL; X17406; CAA35463.1; -.
DR   EMBL; S74659; AAC60643.2; -.
DR   PIR; A39086; A39086.
DR   HSSP; P98066; 1TSG.
DR   Genew; HGNC:319; AGC1.
DR   MIM; 155760; -.
DR   InterPro; IPR006209; EGF_like.
DR   InterPro; IPR006210; IEGF.
```

```
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003599; Ig.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR001304; Lectin_C.
DR   InterPro; IPR000538; Link.
DR   InterPro; IPR003324; SGXXSG.
DR   InterPro; IPR000436; Sushi_SCR_CCP.
DR   Pfam; PF00008; EGF; 1.
DR   Pfam; PF00047; ig; 1.
DR   Pfam; PF00059; lectin_c; 1.
DR   Pfam; PF02339; SGXXSG; 71.
DR   Pfam; PF00084; sushi; 1.
DR   Pfam; PF00193; Xlink; 4.
DR   PRINTS; PR01265; LINKMODULE.
DR   ProDom; PD000918; Link; 4.
DR   SMART; SM00032; CCP; 1.
DR   SMART; SM00034; CLECT; 1.
DR   SMART; SM00181; EGF; 1.
DR   SMART; SM00409; IG; 1.
DR   SMART; SM00445; LINK; 4.
DR   PROSITE; PS00615; C_TYPE_LECTIN_1; 1.
DR   PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
DR   PROSITE; PS00022; EGF_1; 1.
DR   PROSITE; PS01186; EGF_2; 1.
DR   PROSITE; PS50026; EGF_3; 1.
DR   PROSITE; PS50835; IG_LIKE; 1.
DR   PROSITE; PS00290; IG_MHC; 1.
DR   PROSITE; PS01241; LINK; 3.
KW   Glycoprotein; Proteoglycan; Lectin; Signal; Sushi; EGF-like domain;
KW   Alternative splicing; Repeat; Immunoglobulin domain.
FT   SIGNAL        1     19       POTENTIAL.
FT   CHAIN        20   2415       AGGRECAN CORE PROTEIN.
FT   DOMAIN       34    147       IG-LIKE V-TYPE.
FT   DOMAIN      170    247       LINK 1.
FT   DOMAIN      268    349       LINK 2.
FT   DOMAIN      495    572       LINK 3.
FT   DOMAIN      593    673       LINK 4.
FT   DOMAIN     2164   2199       EGF-LIKE.
FT   DOMAIN     2201   2327       C-TYPE LECTIN.
FT   DOMAIN     2331   2389       SUSHI.
FT   DOMAIN       48    141       G1-A.
FT   DOMAIN      152    247       G1-B.
FT   DOMAIN      253    349       G1-B'.
FT   DOMAIN      477    571       G2-B.
FT   DOMAIN      578    672       G2-B'.
FT   DOMAIN      676    848       KS.
FT   DOMAIN      772    844       12 X APPROXIMATE TANDEM REPEATS.
FT   DOMAIN      851   1497       CS-1.
FT   DOMAIN      941   1497       29 X APPROXIMATE TANDEM REPEATS.
FT   DOMAIN     1498   2162       CS-2.
FT   DOMAIN     2163   2415       G3.
FT   DISULFID     51    133       BY SIMILARITY.
FT   DISULFID    175    246       BY SIMILARITY.
FT   DISULFID    199    220       BY SIMILARITY.
FT   DISULFID    273    348       BY SIMILARITY.
FT   DISULFID    297    318       BY SIMILARITY.
FT   DISULFID    500    571       BY SIMILARITY.
FT   DISULFID    524    545       BY SIMILARITY.
FT   DISULFID    598    672       BY SIMILARITY.
FT   DISULFID    621    642       BY SIMILARITY.
FT   DISULFID   2168   2178       BY SIMILARITY.
FT   DISULFID   2173   2187       BY SIMILARITY.
FT   DISULFID   2189   2198       BY SIMILARITY.
FT   DISULFID   2205   2216       BY SIMILARITY.
FT   DISULFID   2233   2325       BY SIMILARITY.
FT   DISULFID   2301   2317       BY SIMILARITY.
FT   DISULFID   2332   2375       BY SIMILARITY.
FT   DISULFID   2361   2388       BY SIMILARITY.
FT   CARBOHYD    126    126       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    239    239       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    333    333       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    387    387       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    434    434       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    602    602       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    657    657       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    737    737       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1898   1898       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   VARSPLIC   2163   2200       Missing (in isoform 2 and isoform 3).
FT                                /FTId=VSP_003074.
FT   VARSPLIC   2330   2390       Missing (In isoform 3).
FT                                /FTId=VSP_003075.
FT   CONFLICT    766    766       E -> A (IN REF. 4).
FT   CONFLICT    847    847       E -> V (IN REF. 4).
FT   CONFLICT   1928   1928       E -> A (IN REF. 2).
FT   CONFLICT   1964   1964       I -> V (IN REF. 2 AND 3).
FT   CONFLICT   2070   2070       P -> A (IN REF. 3).
FT   CONFLICT   2391   2391       A -> P (IN REF. 2 AND 3).
SQ   SEQUENCE   2415 AA;  250191 MW;  1288937E1B98C6B6 CRC64;

  Query Match             14.2%;  Score 158.5;  DB 1;  Length 2415;
  Best Local Similarity   29.1%;  Pred. No. 7.8e-06;
  Matches   52;  Conservative   22;  Mismatches   62;  Indels   43;  Gaps   11;

Qy          9 QPVCRGG---TQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFI 65
              |  |  |   |  ||:  :| |   |  : :|:  ||     | || : :||    :|:
Db       2202 QEVCEEGWNKYQGHCYR--HFPD---RETWVDAERRCREQQSHLSSIVTPEEQ----EFV 2252

Qy         66 ENLLPSDGDF-WIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEPS---CGSEVC 121
                |   :  |: ||||        |    :  : |:||   || ||  ::|       | |
Db       2253 NN---NAQDYQWIGL--------NDRTIEGDFRWSDGHPMQFENWRPNQPDNFFAAGEDC 2301

Qy        122 VVM-YHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYS----DEKPAVPSREAEGEETE 175
              ||| :|:              :|||  ||    | ||       | | |      |:: :
Db       2302 VVMIWHEKG----------EWNDVPCNYHLPFTCKKGTVACGEPPVVEHARTFGQKKD 2349


RESULT 15
PGCA_CANFA
ID   PGCA_CANFA     STANDARD;      PRT;  2333 AA.
AC   Q28343; Q28310;
DT   01-NOV-1997 (Rel. 35, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   Aggrecan core protein precursor (Cartilage-specific proteoglycan core
DE   protein) (CSPCP).
GN   AGC1.
OS   Canis familiaris (Dog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX   NCBI_TaxID=9615;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Glant T.T., Adams M.E., Kwok S.X.F., Huang D., Fueloep C.;
RT   "Complete coding sequence and deduced amino acid sequence of aggrecan
RT   of canine cartilage.";
RL   Submitted (APR-2000) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE OF 774-833 FROM N.A.
RC   TISSUE=Cartilage;
RX   MEDLINE=95128522; PubMed=7827755;
RA   Barry F.P., Neame P.J., Sasse J., Pearson D.;
RT   "Length variation in the keratan sulfate domain of mammalian
RT   aggrecan.";
RL   Matrix Biol. 14:323-328(1994).
RN   [3]
RP   SEQUENCE OF 1830-2333 FROM N.A.
RA   Adams M.E., Kowk S.X.F., Huang D., Glant T.T., Fullop C.;
RL   Submitted (AUG-1996) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   SEQUENCE OF 2082-2118 FROM N.A.
RC   TISSUE=Cartilage;
RX   MEDLINE=93352525; PubMed=8349621;
RA   Fueloep C., Walcz E., Valyon M., Glant T.T.;
RT   "Expression of alternatively spliced epidermal growth factor-like
RT   domains in aggrecans of different species. Evidence for a novel
RT   module.";
```

```
RL   J. Biol. Chem. 268:17377-17383(1993).
CC   -!- FUNCTION: This proteoglycan is a major component of extracellular
CC       matrix of cartilagenous tissues. A major function of this protein
CC       is to resist compression in cartilage. It binds avidly to
CC       hyaluronic acid via an amino-terminal globular region. May play a
CC       regulatory role in the matrix assembly of the cartilage.
CC   -!- SUBCELLULAR LOCATION: Secreted; extracellular matrix (By
CC       similarity).
CC   -!- DOMAIN: Two globular domains, G1 and G2, comprise the amino
CC       terminus of the proteoglycan, while another globular region, G3,
CC       makes up the COOH terminus. G1 contains link domains and thus
CC       consists of three disulfide-bonded loop structures designated as
CC       the A, B, B' motifs. G2 is similar to G1. The keratan sulfate (KS)
CC       and the chondroitin sulfate (CS) attachment domains lie between G2
CC       and G3.
CC   -!- PTM: Contains mostly chondroitin sulfate, but also keratan sulfate
CC       chains, N-linked and O-linked oligosaccharides (By similarity).
CC   -!- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
CC   -!- SIMILARITY: Contains 4 link domains.
CC   -!- SIMILARITY: Contains 1 C-type lectin family domain.
CC   -!- SIMILARITY: Contains 1 Sushi (SCR) domain.
CC   -!- SIMILARITY: Contains 1 EGF-like domain.
CC   -!- SIMILARITY: Belongs to the aggrecan/versican proteoglycan family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U65989; AAB06238.2; -.
DR   EMBL; S74662; AAC60527.1; -.
DR   EMBL; L07054; -; NOT_ANNOTATED_CDS.
DR   PIR; I46998; I46998.
DR   HSSP; P08709; 1BF9.
DR   InterPro; IPR002353; AntifreezeII.
DR   InterPro; IPR000152; Asx_hydroxyl_S.
DR   InterPro; IPR000742; EGF_2.
DR   InterPro; IPR001881; EGF_Ca.
DR   InterPro; IPR006209; EGF_like.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR001304; Lectin_C.
DR   InterPro; IPR000538; Link.
DR   InterPro; IPR003324; SGXXSG.
DR   InterPro; IPR000436; Sushi_SCR_CCP.
DR   Pfam; PF00008; EGF; 1.
DR   Pfam; PF00047; ig; 1.
DR   Pfam; PF00059; lectin_c; 1.
DR   Pfam; PF02339; SGXXSG; 66.
DR   Pfam; PF00084; sushi; 1.
DR   Pfam; PF00193; Xlink; 4.
DR   PRINTS; PR00356; ANTIFREEZEII.
DR   PRINTS; PR01265; LINKMODULE.
DR   ProDom; PD000918; Link; 4.
DR   SMART; SM00032; CCP; 1.
DR   SMART; SM00034; CLECT; 1.
DR   SMART; SM00179; EGF_CA; 1.
DR   SMART; SM00445; LINK; 4.
DR   PROSITE; PS00010; ASX_HYDROXYL; 1.
DR   PROSITE; PS00615; C_TYPE_LECTIN_1; 1.
DR   PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
DR   PROSITE; PS00022; EGF_1; 1.
DR   PROSITE; PS50026; EGF_3; 1.
DR   PROSITE; PS01187; EGF_CA; 1.
DR   PROSITE; PS50835; IG_LIKE; 1.
DR   PROSITE; PS00290; IG_MHC; 1.
DR   PROSITE; PS01241; LINK; 4.
KW   Glycoprotein; Proteoglycan; Lectin; Signal; Sushi; EGF-like domain;
KW   Repeat; Immunoglobulin domain.
FT   SIGNAL        1     16       POTENTIAL.
FT   CHAIN        17   2333       AGGRECAN CORE PROTEIN.
FT   DOMAIN       34    147       IG-LIKE V-TYPE.
FT   DOMAIN      170    247       LINK 1.
FT   DOMAIN      268    349       LINK 2.
FT   DOMAIN      513    590       LINK 3.
FT   DOMAIN      611    692       LINK 4.
FT   DOMAIN     2081   2117       EGF-LIKE, CALCIUM-BINDING (POTENTIAL).
FT   DOMAIN     2130   2245       C-TYPE LECTIN.
FT   DOMAIN     2249   2307       SUSHI.
FT   DOMAIN       48    140       G1-A.
FT   DOMAIN      152    247       G1-B.
FT   DOMAIN      253    349       G1-B'.
FT   DOMAIN      495    589       G2-B.
FT   DOMAIN      596    691       G2-B'.
FT   DOMAIN      694    816       KS.
FT   DOMAIN      819   1394       CS-1.
FT   DOMAIN     1395   2079       CS-2.
FT   DOMAIN     2080   2333       G3.
FT   DISULFID     51    133       BY SIMILARITY.
FT   DISULFID    175    246       BY SIMILARITY.
FT   DISULFID    199    220       BY SIMILARITY.
FT   DISULFID    273    348       BY SIMILARITY.
FT   DISULFID    297    318       BY SIMILARITY.
FT   DISULFID    518    589       BY SIMILARITY.
FT   DISULFID    542    563       BY SIMILARITY.
FT   DISULFID    616    691       BY SIMILARITY.
FT   DISULFID    640    661       BY SIMILARITY.
FT   DISULFID   2085   2096       BY SIMILARITY.
FT   DISULFID   2090   2105       BY SIMILARITY.
FT   DISULFID   2107   2116       BY SIMILARITY.
FT   DISULFID   2123   2134       BY SIMILARITY.
FT   DISULFID   2151   2243       BY SIMILARITY.
FT   DISULFID   2219   2235       BY SIMILARITY.
FT   DISULFID   2250   2293       BY SIMILARITY.
FT   DISULFID   2279   2306       BY SIMILARITY.
FT   CARBOHYD    126    126       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    239    239       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    333    333       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    387    387       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    444    444       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    620    620       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    676    676       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    747    747       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   2333 AA;  240573 MW;  8B9ED78F3508B596 CRC64;

  Query Match           13.9%;  Score 155.5;  DB 1;  Length 2333;
  Best Local Similarity  28.5%;  Pred. No. 1.4e-05;
  Matches   51;  Conservative   24;  Mismatches   61;  Indels   43;  Gaps   11;

Qy          9 QPVCRGG---TQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFI 65
              | :|  |    |  ||:  || |   | :: :|:  ||     | || : :||    :|:
Db       2120 QELCEEGWTKFQGHCYR--YFPD---RESWVDAESRCRAQQSHLSSIVTPEEQ----EFV 2170

Qy         66 ENLLPSDGDF-WIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEPS---CGSEVC 121
                |   :  |: ||||        |    :  : |:||   || ||  ::|       | |
Db       2171 NN---NAQDYQWIGL--------NDRTIEGDFRWSDGHSLQFENWRPNQPDNFFVSGEDC 2219

Qy        122 VVM-YHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYS----DEKPAVPSREAEGEETE 175
              ||| :|:                :|||  ||    | ||        : | |      |:: :
Db       2220 VVMIWHEKG-----------EWNDVPCNYYLPFTCKKGTVACGDPPVVEHARTFGQKKD 2267


RESULT 16
LEM2_MOUSE
ID   LEM2_MOUSE     STANDARD;      PRT;   612 AA.
AC   Q00690;
DT   01-APR-1993 (Rel. 25, Created)
DT   01-APR-1993 (Rel. 25, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   E-selectin precursor (Endothelial leukocyte adhesion molecule 1)
DE   (ELAM-1) (Leukocyte-endothelial cell adhesion molecule 2) (LECAM2).
GN   SELE OR ELAM-1.
OS   Mus musculus (Mouse).
```

```
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=92283265; PubMed=1375914;
RA   Becker-Andre M., van Huijsduijnen R.H., Losberger C., Whelan J.,
RA   Delamarter J.F.;
RT   "Murine endothelial leukocyte-adhesion molecule 1 is a close
RT   structural and functional homologue of the human protein.";
RL   Eur. J. Biochem. 206:401-411(1992).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=92340571; PubMed=1378846;
RA   Weller A., Isenmann S., Vestweber D.;
RT   "Cloning of the mouse endothelial selectins. Expression of both E-
RT   and P-selectin is inducible by tumor necrosis factor alpha.";
RL   J. Biol. Chem. 267:15176-15183(1992).
CC   -!- FUNCTION: Expressed on cytokine induced endothelial cells and
CC       mediates their binding to Leukocytes. The ligand recognized by
CC       ELAM-1 is sialyl-lewis X (alpha(1->3)fucosylated derivatives of
CC       polylactosamine that are found at the nonreducing termini of
CC       glycolipids).
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC   -!- SIMILARITY: Belongs to the selectin/LECAM family.
CC   -!- SIMILARITY: Contains 1 C-type lectin family domain.
CC   -!- SIMILARITY: Contains 1 EGF-like domain.
CC   -!- SIMILARITY: Contains 6 Sushi (SCR) domains.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; M80778; AAA37547.1; -.
DR   EMBL; M87862; AAA37577.1; ALT_INIT.
DR   PIR; S23174; B42755.
DR   HSSP; P16581; 1KJA.
DR   MGD; MGI:98278; Sele.
DR   InterPro; IPR006209; EGF_like.
DR   InterPro; IPR006210; IEGF.
DR   InterPro; IPR001304; Lectin_C.
DR   InterPro; IPR002396; Selectin.
DR   InterPro; IPR000436; Sushi_SCR_CCP.
DR   Pfam; PF00008; EGF; 1.
DR   Pfam; PF00059; lectin_c; 1.
DR   Pfam; PF00084; sushi; 6.
DR   PRINTS; PR00343; SELECTIN.
DR   SMART; SM00032; CCP; 6.
DR   SMART; SM00034; CLECT; 1.
DR   SMART; SM00181; EGF; 1.
DR   PROSITE; PS00615; C_TYPE_LECTIN_1; 1.
DR   PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
DR   PROSITE; PS00022; EGF_1; 1.
DR   PROSITE; PS01186; EGF_2; 1.
DR   PROSITE; PS50026; EGF_3; 1.
KW   Cell adhesion; Transmembrane; Glycoprotein; EGF-like domain; Lectin;
KW   Selectin; Signal; Sushi; Repeat.
FT   SIGNAL       1     21       BY SIMILARITY.
FT   CHAIN       22    612       E-SELECTIN.
FT   DOMAIN      22    557       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   558    579       POTENTIAL.
FT   DOMAIN     580    612       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN      22    139       C-TYPE LECTIN.
FT   DOMAIN     140    175       EGF-LIKE.
FT   DOMAIN     179    239       SUSHI 1.
FT   DOMAIN     242    301       SUSHI 2.
FT   DOMAIN     304    364       SUSHI 3.
FT   DOMAIN     367    427       SUSHI 4.
FT   DOMAIN     430    490       SUSHI 5.
FT   DOMAIN     493    549       SUSHI 6.
FT   DISULFID    40    138       BY SIMILARITY.
FT   DISULFID   111    130       BY SIMILARITY.
FT   DISULFID   143    154       BY SIMILARITY.
FT   DISULFID   148    163       BY SIMILARITY.
FT   DISULFID   165    174       BY SIMILARITY.
FT   DISULFID   180    225       BY SIMILARITY.
FT   DISULFID   210    238       BY SIMILARITY.
FT   DISULFID   243    287       BY SIMILARITY.
FT   DISULFID   273    300       BY SIMILARITY.
FT   DISULFID   305    350       BY SIMILARITY.
FT   DISULFID   336    363       BY SIMILARITY.
FT   DISULFID   368    413       BY SIMILARITY.
FT   DISULFID   399    426       BY SIMILARITY.
FT   DISULFID   431    476       BY SIMILARITY.
FT   DISULFID   462    489       BY SIMILARITY.
FT   DISULFID   494    535       BY SIMILARITY.
FT   DISULFID   521    548       BY SIMILARITY.
FT   CARBOHYD    25     25       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   145    145       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   192    192       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   203    203       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   266    266       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   333    333       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   391    391 .     N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   504    504       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   528    528       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   612 AA;  66749 MW;  86F05713F0EC2C3D CRC64;

  Query Match             13.9%;  Score 154.5;  DB 1;  Length 612;
  Best Local Similarity   27.9%;  Pred. No. 3.3e-06;
  Matches   38;  Conservative   28;  Mismatches   45;  Indels   25;  Gaps    5;

Qy         25 YFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFWIGLRRREE 84
              |:: :|  : ::||   |:||   ||:|:::|   |     ||  |   :|||:|:
Db         23 YYNASSELMTYDEASAYCQRDYTHLVAIQNKEE---INYLNSNLKHSPSYYWIGIRK--- 76

Qy         85 KQSNSTACQDLYAWTDGS---ISQFRNWYVDEPS--CGSEVCVVMYHQPSAPAGIGGPYM 139
                      ::: |         : :||   ||:    :| || :| |  :  :|:
Db         77 -------VNNVWIWVGTGKPLTEEAQNWAPGEPNNKQRNEDCVEIYIQRTKDSGM----- 124

Qy        140 FQWNDDRCNMKNNFIC 155
                |||:||| |   :|
Db        125 --WNDERCNKKKLALC 138


RESULT 17
CD93_MOUSE
ID   CD93_MOUSE     STANDARD;      PRT;   644 AA.
AC   O89103;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Complement component C1q receptor precursor (Complement component 1, q
DE   subcomponent, receptor 1) (C1qRp) (C1qR(p)) (C1q/MBL/SPA receptor)
DE   (CD93 antigen) (Cell surface antigen AA4) (Lymphocyte antigen 68).
GN   C1QR1 OR CD93 OR C1QRP OR LY68 OR AA4.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=129/Sv;
RX   MEDLINE=20528605; PubMed=11074255;
RA   Kim T.S., Park M., Nepomuceno R.R., Palmarini G., Winokur S.,
RA   Cotman C.A., Bengtsson U., Tenner A.J.;
RT   "Characterization of the murine homolog of C1qR(P): identical cellular
RT   expression pattern, chromosomal location and functional activity of
RT   the human and murine C1qR(P).";
RL   Mol. Immunol. 37:377-389(2000).
RN   [2]
RP   SEQUENCE FROM N.A.
```

```
RC   TISSUE=Leukemia;
RX   MEDLINE=99330438; PubMed=10403644;
RA   Petrenko O., Beavis A., Klaine M., Kittappa R., Godin I.,
RA   Lemischka I.R.;
RT   "The molecular characterization of the fetal stem cell marker AA4.";
RL   Immunity 10:691-700(1999).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=129/Sv; TISSUE=Endothelial cells, and Spleen;
RX   MEDLINE=99359842; PubMed=10430665;
RA   Norsworthy P.J., Taylor P.R., Walport M.J., Botto M.;
RT   "Cloning of the mouse homolog of the 126-kDa human C1q/MBL/SP-A
RT   receptor, C1qRp.";
RL   Mamm. Genome 10:789-793(1999).
CC   -!- FUNCTION: Receptor (or element of a larger receptor complex) for
CC       C1q, mannose-binding lectin (MBL2) and pulmonary surfactant
CC       protein A (SPA). May mediate the enhancement of phagocytosis in
CC       monocytes and macrophages upon interaction with soluble defense
CC       collagens. May play a role in intercellular adhesion. Marker for
CC       early multipotent hematopoeitic precursor cells. May play a role
CC       in cell-cell interactions during hematopoietic and vascular
CC       development.
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC   -!- TISSUE SPECIFICITY: Expressed in lung, heart and bone marrow.
CC       Expressed at lower level in ovary, whole embryo and fetal liver.
CC       Not detected in brain, adult liver or thymus. Highly expressed in
CC       peritoneal cavity and bone marrow macrophages. Not detected in
CC       epithelial cells.
CC   -!- DEVELOPMENTAL STAGE: First detectable in day 9 embryos, in the
CC       endocardium and vascular endothelium in the anterior part of the
CC       embryo. Expression in endothelial cells, initially restricted to
CC       aorta, omphalomesenteric and umbilical arteries, later extends to
CC       subcardinal veins, intersomitic arteries and perimeural vessels.
CC       On day 10, detectable in the entire embryo.
CC   -!- PTM: N- and O-glycosylated (By similarity).
CC   -!- SIMILARITY: Contains 1 C-type lectin family domain.
CC   -!- SIMILARITY: Contains 5 EGF-like domains.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF074856; AAC63274.1; -.
DR   EMBL; AF081789; AAC62649.1; -.
DR   EMBL; AF099939; AAD47906.1; -.
DR   EMBL; AF099938; AAD47906.1; JOINED.
DR   HSSP; P35555; 1EMN.
DR   MGD; MGI:106664; C1qr1.
DR   GO; GO:0016023; C:cytoplasmic vesicle; IDA.
DR   GO; GO:0016021; C:integral to membrane; ISS.
DR   GO; GO:0005886; C:plasma membrane; IDA.
DR   GO; GO:0004872; F:receptor activity; ISS.
DR   GO; GO:0016337; P:cell-cell adhesion; ISS.
DR   GO; GO:0042116; P:macrophage activation; ISS.
DR   GO; GO:0006909; P:phagocytosis; ISS.
DR   InterPro; IPR000152; Asx_hydroxyl_S.
DR   InterPro; IPR001881; EGF_Ca.
DR   InterPro; IPR006209; EGF_like.
DR   InterPro; IPR001304; Lectin_C.
DR   Pfam; PF00008; EGF; 5.
DR   Pfam; PF00059; lectin_c; 1.
DR   SMART; SM00034; CLECT; 1.
DR   SMART; SM00179; EGF_CA; 3.
DR   PROSITE; PS00010; ASX_HYDROXYL; 3.
DR   PROSITE; PS00615; C_TYPE_LECTIN_1; FALSE_NEG.
DR   PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
DR   PROSITE; PS01186; EGF_2; 3.
DR   PROSITE; PS50026; EGF_3; 4.
DR   PROSITE; PS01187; EGF_CA; 3.
KW   Cell adhesion; Receptor; Repeat; Signal; Transmembrane;
KW   EGF-like domain; Lectin; Glycoprotein.
FT   SIGNAL        1    22       POTENTIAL.
FT   CHAIN        23   644       COMPLEMENT COMPONENT C1Q RECEPTOR.
FT   DOMAIN       23   572       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    573   593       POTENTIAL.
FT   DOMAIN      594   644       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN       31   173       C-TYPE LECTIN.
FT   DOMAIN      257   298       EGF-LIKE 1.
FT   DOMAIN      299   341       EGF-LIKE 2.
FT   DOMAIN      342   381       EGF-LIKE 3, CALCIUM-BINDING (POTENTIAL).
FT   DOMAIN      382   423       EGF-LIKE 4, CALCIUM-BINDING (POTENTIAL).
FT   DOMAIN      424   465       EGF-LIKE 5, CALCIUM-BINDING (POTENTIAL).
FT   DISULFID    261   272       BY SIMILARITY.
FT   DISULFID    268   282       BY SIMILARITY.
FT   DISULFID    284   297       BY SIMILARITY.
FT   DISULFID    303   314       BY SIMILARITY.
FT   DISULFID    308   325       BY SIMILARITY.
FT   DISULFID    327   340       BY SIMILARITY.
FT   DISULFID    346   355       BY SIMILARITY.
FT   DISULFID    351   364       BY SIMILARITY.
FT   DISULFID    366   380       BY SIMILARITY.
FT   DISULFID    386   397       BY SIMILARITY.
FT   DISULFID    393   406       BY SIMILARITY.
FT   DISULFID    408   422       BY SIMILARITY.
FT   DISULFID    428   440       BY SIMILARITY.
FT   DISULFID    436   449       BY SIMILARITY.
FT   DISULFID    451   464       BY SIMILARITY.
FT   CARBOHYD    102   102       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    322   322       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   644 AA;  69354 MW;  EB4351648BF8635A CRC64;

  Query Match             13.8%;  Score 153.5;  DB 1;  Length 644;
  Best Local Similarity   24.1%;  Pred. No. 4.3e-06;
  Matches   51;  Conservative   32;  Mismatches   70;  Indels   59;  Gaps   11;

Qy          6 LSGQP-----------VCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIES 54
              | |||           || |     ||   :     :|:  ||:  |  :|| | :::|
Db         13 LLGQPWAGAAADSQAVVCEG---TACYTAHW-----GKLSAAEAQHRCNENGGNLATVKS 64

Qy         55 EDEQKLIEKFIENLLPSD-------GDFWIGLRRREEKQSNSTACQDL----YAWT-DGS 102
              |:| : ::: :  || :       | |||||:|    :  :    ||     ::|   |
Db         65 EEEARHVQQALTQLLKTKAPLEAKMGKFWIGLQR----EKGNCTYHDLPMRGFSWVGGGE 120

Qy        103 ISQFRNWY-VDEPSCGSEVCVVMY-------HQPSAPAGIGGPYMFQWNDDRC------- 147
               : : |||   : ||  : || :         |    |          :|:: |
Db        121 DTAYSNWYKASKSSCIFKRCVSLILDLSLTPHPSHLP---------KWHESPCGTPEAPG 171

Qy        148 NMKNNFICKYSDEKPAVPSREAEGEETELTTP 179
              |    |:||:: :    |          |||
Db        172 NSIEGFLCKFNFKGMCRPLALGGPGRVTYTTP 203


RESULT 18
PGCA_RAT
ID   PGCA_RAT       STANDARD;      PRT;  2124 AA.
AC   P07897;
DT   01-AUG-1988 (Rel. 08, Created)
DT   01-FEB-1991 (Rel. 17, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   Aggrecan core protein precursor (Cartilage-specific proteoglycan core
DE   protein) (CSPCP).
GN   AGC1 OR AGC.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=88087070; PubMed=3693370;
RA   Doege K., Sasaki M., Horigan E., Hassell J.R., Yamada Y.;
RT   "Complete primary structure of the rat cartilage proteoglycan core
RT   protein deduced from cDNA clones.";
```

```
RL   J. Biol. Chem. 262:17757-17767(1987).
RN   [2]
RP   REVISION TO 698.
RA   Doege K., Sasaki M., Horigan E., Hassell J.R., Yamada Y.;
RL   J. Biol. Chem. 263:10040-10040(1988).
RN   [3]
RP   SEQUENCE OF 1856-2124 FROM N.A.
RX   MEDLINE=86250698; PubMed=2424893;
RA   Doege K., Fernandez P., Hassell J.R., Sasaki M., Yamada Y.;
RT   "Partial cDNA sequence encoding a globular domain at the C terminus
RT   of the rat cartilage proteoglycan.";
RL   J. Biol. Chem. 261:8108-8111(1986).
CC   -!- FUNCTION: This proteoglycan is a major component of extracellular
CC       matrix of cartilagenous tissues. A major function of this protein
CC       is to resist compression in cartilage. It binds avidly to
CC       hyaluronic acid via an amino-terminal globular region. May play a
CC       regulatory role in the matrix assembly of the cartilage.
CC   -!- SUBUNIT: Interacts with FBLN1 (By similarity).
CC   -!- SUBCELLULAR LOCATION: Secreted; extracellular matrix (By
CC       similarity).
CC   -!- DOMAIN: Two globular domains, G1 and G2, comprise the amino
CC       terminus of the proteoglycan, while another globular region, G3,
CC       makes up the COOH terminus. G1 contains link domains and thus
CC       consists of three disulfide-bonded loop structures designated as
CC       the A, B, B' motifs. G2 is similar to G1. The keratan sulfate (KS)
CC       and the chondroitin sulfate (CS) attachment domains lie between G2
CC       and G3.
CC   -!- PTM: Contains mostly chondroitin sulfate, but also keratan sulfate
CC       chains, N-linked and O-linked oligosaccharides.
CC   -!- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
CC   -!- SIMILARITY: Contains 4 link domains.
CC   -!- SIMILARITY: Contains 1 C-type lectin family domain.
CC   -!- SIMILARITY: Contains 1 Sushi (SCR) domain.
CC   -!- SIMILARITY: Belongs to the aggrecan/versican proteoglycan family.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; M13518; AAA41836.1; -.
DR   EMBL; J03485; AAA21000.1; ALT_SEQ.
DR   PIR; A92623; A28452.
DR   HSSP; P98066; 1TSG.
DR   InterPro; IPR002353; AntifreezeII.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR003596; Ig_v.
DR   InterPro; IPR001304; Lectin_C.
DR   InterPro; IPR000538; Link.
DR   InterPro; IPR003324; SGXXSG.
DR   InterPro; IPR000436; Sushi_SCR_CCP.
DR   Pfam; PF00047; ig; 1.
DR   Pfam; PF00059; lectin_c; 1.
DR   Pfam; PF02339; SGXXSG; 55.
DR   Pfam; PF00084; sushi; 1.
DR   Pfam; PF00193; Xlink; 4.
DR   PRINTS; PR00356; ANTIFREEZEII.
DR   PRINTS; PR01265; LINKMODULE.
DR   ProDom; PD000918; Link; 4.
DR   SMART; SM00032; CCP; 1.
DR   SMART; SM00034; CLECT; 1.
DR   SMART; SM00406; IGv; 1.
DR   SMART; SM00445; LINK; 4.
DR   PROSITE; PS00615; C_TYPE_LECTIN_1; 1.
DR   PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
DR   PROSITE; PS50835; IG_LIKE; 1.
DR   PROSITE; PS00290; IG_MHC; 1.
DR   PROSITE; PS01241; LINK; 4.
KW   Glycoprotein; Proteoglycan; Lectin; Signal; Sushi; Repeat;
KW   Immunoglobulin domain.
FT   SIGNAL        1     19       POTENTIAL.
FT   CHAIN        20   2124       AGGRECAN CORE PROTEIN.
FT   DOMAIN       34    147       IG-LIKE V-TYPE.
FT   DOMAIN      170    247       LINK 1.
FT   DOMAIN      268    349       LINK 2.
FT   DOMAIN      504    581       LINK 3.
FT   DOMAIN      602    683       LINK 4.
FT   DOMAIN     1910   2036       C-TYPE LECTIN.
FT   DOMAIN     2040   2098       SUSHI.
FT   DOMAIN       48    140       G1-A.
FT   DOMAIN      152    247       G1-B.
FT   DOMAIN      253    349       G1-B'.
FT   DOMAIN      486    580       G2-B.
FT   DOMAIN      587    682       G2-B'.
FT   DOMAIN      685    798       KS.
FT   DOMAIN      801   1226       CS-1.
FT   DOMAIN     1227   1909       CS-2.
FT   DOMAIN     1910   2124       G3.
FT   DISULFID     51    133       BY SIMILARITY.
FT   DISULFID    175    246       BY SIMILARITY.
FT   DISULFID    199    220       BY SIMILARITY.
FT   DISULFID    273    348       BY SIMILARITY.
FT   DISULFID    297    318       BY SIMILARITY.
FT   DISULFID    509    580       BY SIMILARITY.
FT   DISULFID    533    554       BY SIMILARITY.
FT   DISULFID    607    682       BY SIMILARITY.
FT   DISULFID    631    652       BY SIMILARITY.
FT   DISULFID   1914   1925       BY SIMILARITY.
FT   DISULFID   1942   2034       BY SIMILARITY.
FT   DISULFID   2010   2026       BY SIMILARITY.
FT   DISULFID   2041   2084       BY SIMILARITY.
FT   DISULFID   2070   2097       BY SIMILARITY.
FT   CARBOHYD    126    126       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    239    239       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    333    333       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    387    387       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    611    611       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    667    667       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1842   1842       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   2124 AA;  221117 MW;  E30BBE61593A34B1 CRC64;

  Query Match             13.8%;  Score 153.5;  DB 1;  Length 2124;
  Best Local Similarity   26.3%;  Pred. No. 1.9e-05;
  Matches   47;  Conservative   27;  Mismatches   62;  Indels   43;  Gaps    9;

Qy          9 QPVCRGG---TQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFI 65
              |  |  |    |  ||:  :| |   |  : :|:  ||     | || : :||: : |
Db       1911 QEQCEEGWTKFQGHCYR--HFPD---RETWVDAERRCREQQSHLSSIVTPEEQEFVNKNA 1965

Qy         66 ENLLPSDGDF-WIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEP----SCGSEV 120
              :       |: ||||        |    :  : |:||   ||  |  ::|    : | :
Db       1966 Q-------DYQWIGL--------NDRTIEGDFRWSDGHSLQFEKWRPNQPDNFFATGEDC 2010

Qy        121 CVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYS----DEKPAVPSREAEGEETE 175
               |:::|:           :|||  || :  | ||       | |||      |:: :
Db       2011 VVMIWHERG----------EWNDVPCNYQLPFTCKKGTVACGEPPAVEHARTLGQKKD 2058


RESULT 19
PGCB_BOVIN
ID   PGCB_BOVIN     STANDARD;      PRT;   912 AA.
AC   Q28062;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   Brevican core protein precursor.
GN   BCAN.
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
OX   NCBI_TaxID=9913;
```

```
RN   [1]
RP   SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC   TISSUE=Brain;
RX   MEDLINE=94193597; PubMed=8144512;
RA   Yamada H., Watanabe K., Shimonaka M., Yamaguchi Y.;
RT   "Molecular cloning of brevican, a novel brain proteoglycan of the
RT   aggrecan/versican family.";
RL   J. Biol. Chem. 269:10119-10126(1994).
CC   -!- FUNCTION: May play a role in the terminally differentiating and
CC       the adult nervous system during postnatal development. Could
CC       stabilize interactions between HA and brain proteoglycans.
CC   -!- SUBCELLULAR LOCATION: Secreted; extracellular matrix (By
CC       similarity).
CC   -!- TISSUE SPECIFICITY: Brain; expressed in cerebellar astrocytes but
CC       not in neurons.
CC   -!- PTM: Contains mostly chondroitin sulfate.
CC   -!- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
CC   -!- SIMILARITY: Contains 2 link domains.
CC   -!- SIMILARITY: Contains 1 EGF-like domain.
CC   -!- SIMILARITY: Contains 1 C-type lectin family domain.
CC   -!- SIMILARITY: Contains 1 Sushi (SCR) domain.
CC   -!- SIMILARITY: Belongs to the aggrecan/versican proteoglycan family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X75887; CAA53481.1; -.
DR   PIR; A54423; A54423.
DR   HSSP; P20693; 1HLJ.
DR   InterPro; IPR002353; AntifreezeII.
DR   InterPro; IPR000742; EGF_2.
DR   InterPro; IPR006209; EGF_like.
DR   InterPro; IPR006210; IEGF.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR003596; Ig_v.
DR   InterPro; IPR001304; Lectin_C.
DR   InterPro; IPR000538; Link.
DR   InterPro; IPR000436; Sushi_SCR_CCP.
DR   Pfam; PF00008; EGF; 1.
DR   Pfam; PF00047; ig; 1.
DR   Pfam; PF00059; lectin_c; 1.
DR   Pfam; PF00084; sushi; 1.
DR   Pfam; PF00193; Xlink; 2.
DR   PRINTS; PR00356; ANTIFREEZEII.
DR   PRINTS; PR01265; LINKMODULE.
DR   ProDom; PD000918; Link; 2.
DR   SMART; SM00032; CCP; 1.
DR   SMART; SM00034; CLECT; 1.
DR   SMART; SM00181; EGF; 1.
DR   SMART; SM00406; IGv; 1.
DR   SMART; SM00445; LINK; 2.
DR   PROSITE; PS00615; C_TYPE_LECTIN_1; 1.
DR   PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
DR   PROSITE; PS00022; EGF_1; 1.
DR   PROSITE; PS01186; EGF_2; 1.
DR   PROSITE; PS50026; EGF_3; 1.
DR   PROSITE; PS50835; IG_LIKE; 1.
DR   PROSITE; PS00290; IG_MHC; 1.
DR   PROSITE; PS01241; LINK; 2.
KW   Glycoprotein; Hyaluronic acid; Proteoglycan; Lectin; Signal; Sushi;
KW   EGF-like domain; Repeat; Immunoglobulin domain.
FT   SIGNAL       1     22       POTENTIAL.
FT   CHAIN       23    912       BREVICAN CORE PROTEIN.
FT   DOMAIN      36    155       IG-LIKE V-TYPE.
FT   DOMAIN     174    251       LINK 1.
FT   DOMAIN     272    353       LINK 2.
FT   DOMAIN     647    683       EGF-LIKE.
FT   DOMAIN     683    811       C-TYPE LECTIN.
FT   DOMAIN     812    876       SUSHI.
FT   DISULFID    57    137       BY SIMILARITY.
FT   DISULFID   179    250       BY SIMILARITY.
FT   DISULFID   203    224       BY SIMILARITY.
FT   DISULFID   277    352       BY SIMILARITY.
FT   DISULFID   301    322       BY SIMILARITY.
FT   DISULFID   651    662       BY SIMILARITY.
FT   DISULFID   656    671       BY SIMILARITY.
FT   DISULFID   673    682       BY SIMILARITY.
FT   DISULFID   689    700       BY SIMILARITY.
FT   DISULFID   717    809       BY SIMILARITY.
FT   DISULFID   785    801       BY SIMILARITY.
FT   DISULFID   816    859       BY SIMILARITY.
FT   DISULFID   845    872       BY SIMILARITY.
FT   CARBOHYD   130    130       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   337    337       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   912 AA;  99554 MW;  677B3EB1C688C4D7 CRC64;

  Query Match             13.6%;  Score 152;  DB 1;  Length 912;
  Best Local Similarity   28.8%;  Pred. No. 8.9e-06;
  Matches   47;  Conservative   17;  Mismatches   61;  Indels   38;  Gaps    8;

Qy         17 QRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFW 76
              |  |||    | ::|| ::|||:  ||  |  | || : :||  |           |
Db        697 QGACYK----HFSARR-SWEEAENKCRMYGAHLASISTPEEQDFINNRYREYQ------W 745

Qy         77 IGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEPS---CGSEVCVVM-YHQPSAPA 132
              |||         |    :  : |:||    : ||   :|       | |||| :|
Db        746 IGL--------NDRTIEGDFLWSDGVPLLYENWNPGQPDSYFLSGENCVVMVWHDQG--- 794

Qy        133 GIGGPYMFQWNDDRCNMKNNFICKYS----DEKPAVPSREAEG 171
                     ||:|  ||   :: ||          | :|  |  |
Db        795 --------QWSDVPCNYHLSYTCKMGLVSCGPPPELPLAEVFG 829


RESULT 20
CD93_HUMAN
ID   CD93_HUMAN     STANDARD;      PRT;   652 AA.
AC   Q9NPY3; O00274;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Complement component C1q receptor precursor (Complement component 1, q
DE   subcomponent, receptor 1) (C1qRp) (C1qR(p)) (C1q/MBL/SPA receptor)
DE   (CD93 antigen) (CDw93).
GN   C1QR1 OR CD93.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX   MEDLINE=97199258; PubMed=9047234;
RA   Nepomuceno R.R., Henschen-Edman A.H., Burgess W.H., Tenner A.J.;
RT   "cDNA cloning and primary structure analysis of C1qR(P), the human
RT   C1q/MBL/SPA receptor that mediates enhanced phagocytosis in vitro.";
RL   Immunity 6:119-129(1997).
RN   [2]
RP   SEQUENCE FROM N.A., AND VARIANT ALA-318..
RX   MEDLINE=21640567; PubMed=11781389;
RA   Steinberger P., Szekeres A., Wille S., Stockl J., Selenko N.,
RA   Prager E., Staffler G., Madic O., Stockinger H., Knapp W.;
RT   "Identification of human CD93 as the phagocytic C1q receptor (C1qRp)
RT   by expression cloning.";
RL   J. Leukoc. Biol. 71:133-140(2002).
RN   [3]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21638749; PubMed=11780052;
RA   Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,
RA   Jones M., Stavrides G., Almeida J.P., Babbage A.K., Bagguley C.L.,
RA   Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,
RA   Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,
```

```
RA   Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,
RA   Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,
RA   Clegg S., Cobley V.E., Collier R.E., Connor R.E., Corby N.R.,
RA   Coulson A., Coville G.J., Deadman R., Dhami P.D., Dunn M.,
RA   Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,
RA   Grafham D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,
RA   Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,
RA   Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,
RA   Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,
RA   Lehvaeslaiho M.H., Leversha M.A., Lloyd C., Lloyd D.M., Lovell J.D.,
RA   Marsh V.L., Martin S.L., McConnachie L.J., McLay K., McMurray A.A.,
RA   Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,
RA   Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,
RA   Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsay H.,
RA   Rice C.M., Ross M.T., Scott C.E., Sehra H.K., Shownkeen R., Sims S.,
RA   Skuce C.D., Smith M.L., Soderlund C., Steward C.A., Sulston J.E.,
RA   Swann R.M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,
RA   Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M.,
RA   Whitehead S.L., Whittaker P., Willey D.L., Williams L., Williams S.A.,
RA   Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,
RA   Rogers J.;
RT   "The DNA sequence and comparative analysis of human chromosome 20.";
RL   Nature 414:865-871(2001).
RN   [4]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Leukocyte;
RX   MEDLINE=22388257; PubMed=12477932;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length
RT   human and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [5]
RP   CHARACTERIZATION.
RX   MEDLINE=21990337; PubMed=11994479;
RA   McGreal E.P., Ikewaki N., Akatsu H., Morgan B.P., Gasque P.;
RT   "Human C1qRp is identical with CD93 and the mNI-11 antigen but does
RT   not bind C1q.";
RL   J. Immunol. 168:5222-5232(2002).
RN   [6]
RP   O-GLYCOSYLATION.
RX   MEDLINE=99192777; PubMed=10092817;
RA   Nepomuceno R.R., Ruiz S., Park M., Tenner A.J.;
RT   "C1qRP is a heavily O-glycosylated cell surface protein involved in
RT   the regulation of phagocytic activity.";
RL   J. Immunol. 162:3583-3589(1999).
CC   -!- FUNCTION: Receptor (or element of a larger receptor complex) for
CC       C1q, mannose-binding lectin (MBL2) and pulmonary surfactant
CC       protein A (SPA). May mediate the enhancement of phagocytosis in
CC       monocytes and macrophages upon interaction with soluble defense
CC       collagens. May play a role in intercellular adhesion.
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC   -!- TISSUE SPECIFICITY: Highly expressed in endothelial cells,
CC       platelets, cells of myeloid origin, such as monocytes and
CC       neutrophils. Not expressed in cells of lymphoid origin.
CC   -!- PTM: N- and O-glycosylated.
CC   -!- SIMILARITY: Contains 1 C-type lectin family domain.
CC   -!- SIMILARITY: Contains 5 EGF-like domains.
CC   -!- CAUTION: Has been sometimes referred to as a collectin receptor.
CC   -!- CAUTION: According to Ref.5, C1q is not a ligand for C1QR1.
CC   -!- DATABASE: NAME=PROW; NOTE=PROW 3:1-6(2001);
CC       WWW="http://www.ncbi.nlm.nih.gov/prow/guide/467246456_g.htm".
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U94333; AAB53110.1; -.
DR   EMBL; AL118508; CAC00597.1; -.
DR   EMBL; BC028075; AAH28075.1; -.
DR   HSSP; P35555; 1EMN.
DR   Genew; HGNC:15855; C1QR1.
DR   MIM; 120577; -.
DR   GO; GO:0016021; C:integral to membrane; IC.
DR   GO; GO:0004872; F:receptor activity; NAS.
DR   GO; GO:0016337; P:cell-cell adhesion; IDA.
DR   GO; GO:0042116; P:macrophage activation; NAS.
DR   GO; GO:0006909; P:phagocytosis; NAS.
DR   InterPro; IPR000152; Asx_hydroxyl_S.
DR   InterPro; IPR001881; EGF_Ca.
DR   InterPro; IPR006209; EGF_like.
DR   InterPro; IPR001304; Lectin_C.
DR   Pfam; PF00008; EGF; 5.
DR   Pfam; PF00059; lectin_c; 1.
DR   SMART; SM00034; CLECT; 1.
DR   SMART; SM00179; EGF_CA; 3.
DR   PROSITE; PS00010; ASX_HYDROXYL; 3.
DR   PROSITE; PS00615; C_TYPE_LECTIN_1; FALSE_NEG.
DR   PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
DR   PROSITE; PS01186; EGF_2; 3.
DR   PROSITE; PS50026; EGF_3; 3.
DR   PROSITE; PS01187; EGF_CA; 3.
KW   Cell adhesion; Receptor; Repeat; Signal; Transmembrane;
KW   EGF-like domain; Lectin; Glycoprotein; Polymorphism.
FT   SIGNAL        1     21
FT   CHAIN        22    652       COMPLEMENT COMPONENT C1Q RECEPTOR.
FT   DOMAIN       24    580       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    581    601       POTENTIAL.
FT   DOMAIN      602    652       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN       32    174       C-TYPE LECTIN.
FT   DOMAIN      260    301       EGF-LIKE 1.
FT   DOMAIN      302    344       EGF-LIKE 2.
FT   DOMAIN      345    384       EGF-LIKE 3, CALCIUM-BINDING (POTENTIAL).
FT   DOMAIN      385    426       EGF-LIKE 4, CALCIUM-BINDING (POTENTIAL).
FT   DOMAIN      427    468       EGF-LIKE 5, CALCIUM-BINDING (POTENTIAL).
FT   DOMAIN      594    601       POLY-LEU.
FT   DISULFID    264    275       BY SIMILARITY.
FT   DISULFID    271    285       BY SIMILARITY.
FT   DISULFID    287    300       BY SIMILARITY.
FT   DISULFID    306    317       BY SIMILARITY.
FT   DISULFID    311    328       BY SIMILARITY.
FT   DISULFID    330    343       BY SIMILARITY.
FT   DISULFID    349    358       BY SIMILARITY.
FT   DISULFID    354    367       BY SIMILARITY.
FT   DISULFID    369    383       BY SIMILARITY.
FT   DISULFID    389    400       BY SIMILARITY.
FT   DISULFID    396    409       BY SIMILARITY.
FT   DISULFID    411    425       BY SIMILARITY.
FT   DISULFID    431    443       BY SIMILARITY.
FT   DISULFID    439    452       BY SIMILARITY.
FT   DISULFID    454    467       BY SIMILARITY.
FT   CARBOHYD    325    325       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   VARIANT     318    318       V -> A.
FT                                 /FTId=VAR_013573.
FT   CONFLICT     22     22       T -> V (IN REF. 1; AA SEQUENCE).
FT   CONFLICT     36     36       C -> T (IN REF. 1; AA SEQUENCE).
FT   CONFLICT     38     39       TA -> RI (IN REF. 1; AA SEQUENCE).
```

```
FT   CONFLICT    155    155       S -> N (IN REF. 1).
FT   CONFLICT    186    186       G -> A (IN REF. 1; AA SEQUENCE).
FT   CONFLICT    492    492       S -> A (IN REF. 1; AA SEQUENCE).
FT   CONFLICT    496    496       R -> Q (IN REF. 1; AA SEQUENCE).
FT   CONFLICT    504    504       R -> G (IN REF. 1; AA SEQUENCE).
FT   CONFLICT    541    541       P -> S (IN REF. 1).
SQ   SEQUENCE   652 AA;  68560 MW;  EECA0FEAC55FCAC2 CRC64;

  Query Match             13.6%;  Score 151.5;  DB 1;  Length 652;
  Best Local Similarity   24.4%;  Pred. No. 6.6e-06;
  Matches   50;  Conservative   38;  Mismatches   74;  Indels   43;  Gaps    9;

Qy          5 LLSGQP-----------VCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIE 53
              ||  ||             || |      ||        | :|:  ||:  | ::|| | :::
Db         13 LLLTQPGAGTGADTEAVVCVG---TACYTA-----HSGKLSAAEAQNHCNQNGGNLATVK 64

Qy         54 SEDEQKLIEKFIENLLPSD-------GDFWIGLRRREEKQSNSTACQDLYAWT-DGSISQ 105
              |::| : ::: :  ||  :          |||||:| : |  : :     ::|   |  :
Db         65 SKEEAQHVQRVLAQLLRREAALTARMSKFWIGLQREKGKCLDPSLPLKGFSWVGGGEDTP 124

Qy        106 FRNWYVD-EPSCGSEVCVVM---YHQPSAPAGIGGPYMFQWNDDRCNMKNN-------FI 154
              : ||: :   || |: || :     ||  |: :      :|:: |   :        |:
Db        125 YSNWHKELRNSCISKRCVSLLLDLSQPLLPSRLP-----KWSEGPCGSPGSPGSNIEGFV 179

Qy        155 CKYSDEKPAVPSREAEGEETELTTP 179
              ||:| :    |        :  |||
Db        180 CKFSFKGMCRPLALGGPGQVTYTTP 204


RESULT 21
CLE1_HUMAN
ID   CLE1_HUMAN     STANDARD;      PRT;   197 AA.
AC   O75596;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   C-type lectin superfamily member 1 precursor (Cartilage-derived C-type
DE   lectin).
GN   CLECSF1.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Cartilage;
RX   MEDLINE=99453719; PubMed=10524194;
RA   Neame P.J., Tapp H., Grimm D.R.;
RT   "The cartilage-derived, C-type lectin (CLECSF1): structure of the gene
RT   and chromosomal location.";
RL   Biochim. Biophys. Acta 1446:193-202(1999).
CC   -!- TISSUE SPECIFICITY: Restricted to cartilage.
CC   -!- SIMILARITY: Contains 1 C-type lectin family domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF077345; AAD12542.1; -.
DR   EMBL; AF077344; AAD12542.1; JOINED.
DR   HSSP; P05452; 1HTN.
DR   Genew; HGNC:2052; CLECSF1.
DR   GO; GO:0005530; F:lectin; TAS.
DR   GO; GO:0001501; P:skeletal development; TAS.
DR   InterPro; IPR001304; Lectin_C.
DR   Pfam; PF00059; lectin_c; 1.
DR   SMART; SM00034; CLECT; 1.
DR   PROSITE; PS00615; C_TYPE_LECTIN_1; 1.
DR   PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
KW   Lectin; Signal.
FT   SIGNAL        1     22       POTENTIAL.
FT   CHAIN        23    197       C-TYPE LECTIN SUPERFAMILY MEMBER 1.
FT   DOMAIN       74    192       C-TYPE LECTIN.
FT   DISULFID     68     78       BY SIMILARITY.
FT   DISULFID     95    191       BY SIMILARITY.
FT   DISULFID    167    183       BY SIMILARITY.
SQ   SEQUENCE   197 AA;  22232 MW;  BB924DBDDB7729A4 CRC64;

  Query Match             13.5%;  Score 151;  DB 1;  Length 197;
  Best Local Similarity   25.0%;  Pred. No. 1.7e-06;
  Matches   38;  Conservative   26;  Mismatches   64;  Indels   24;  Gaps    4;

Qy          9 QPVCRGGTQ--RPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIE 66
              | ||  ||:  : ||       :    :| || | |   || ||   : ||   :: : :
Db         65 QTVCLRGTKVHKKCYLA-----SEGLKHFHEANEDCISKGGILVIPRNSDEINALQDYGK 119

Qy         67 NLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEPSCGSEVCVVMYH 126
              ||   |||:|:         |    :  :   :|    | ||   :|: |      |::
Db        120 RSLPGVNDFWLGI--------NDMVTEGKFVDVNGIAISFLNWDRAQPNGGKRENCVLFS 171

Qy        127 QPSAPAGIGGPYMFQWNDDRCNMKNNFICKYS 158
              | :           :|:|: |     :||:::
Db        172 QSA---------QGKWSDEACRSSKRYICEFT 194


RESULT 22
FCE2_HUMAN
ID   FCE2_HUMAN     STANDARD;      PRT;   321 AA.
AC   P06734;
DT   01-JAN-1988 (Rel. 06, Created)
DT   01-JAN-1988 (Rel. 06, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   Low affinity immunoglobulin epsilon FC receptor (Lymphocyte IgE
DE   receptor) (FC-epsilon-RII) (CD23) (BLAST-2) (Immunoglobulin E-binding
DE   factor).
GN   FCER2 OR IGEBF.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=87118255; PubMed=2949326;
RA   Ikuta K., Takami M., Kim C.W., Honjo T., Miyoshi T., Tagaya Y.,
RA   Kawabe T., Yodoi J.;
RT   "Human lymphocyte Fc receptor for IgE: sequence homology of its
RT   cloned cDNA with animal lectins.";
RL   Proc. Natl. Acad. Sci. U.S.A. 84:819-823(1987).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=87051737; PubMed=2877743;
RA   Kikutani H., Inui S., Sato R., Barsumian E.L., Owaki H.,
RA   Yamasaki K., Kaisho T., Uchibayashi N., Hardy R.R., Hirano T.,
RA   Tsunasawa S., Sakiyama F., Suemura M., Kishimoto T.;
RT   "Molecular structure of human lymphocyte receptor for immunoglobulin
RT   E.";
RL   Cell 47:657-665(1986).
RN   [3]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=87218454; PubMed=3034567;
RA   Luedin C., Hofstetter H., Sarfati M., Levy C.A., Suter U., Alaimo D.,
RA   Kilchherr E., Frost H., Delespesse G.;
RT   "Cloning and expression of the cDNA coding for a human lymphocyte IgE
RT   receptor.";
RL   EMBO J. 6:109-114(1987).
RN   [4]
RP   PARTIAL SEQUENCE, AND DISULFIDE BONDS.
RX   MEDLINE=93038513; PubMed=1417742;
RA   Rose K., Turcatti G., Graber P., Pochon S., Regamey P.-O.,
RA   Jansen K.U., Magnenat E., Aubonney N., Bonnefoy J.-Y.;
RT   "Partial characterization of natural and recombinant human soluble
RT   CD23.";
```

```
RL   Biochem. J. 286:819-824(1992).
RN   [5]
RP   ALTERNATIVE SPLICING.
RX   MEDLINE=89028672; PubMed=2972386;
RA   Yokota A., Kikutani H., Tanaka T., Sato R., Barsumian E.L.,
RA   Suemura M., Kishimoto T.;
RT   "Two species of human Fc epsilon receptor II (Fc epsilon RII/CD23):
RT   tissue-specific and IL-4-specific regulation of gene expression.";
RL   Cell 55:611-618(1988).
RN   [6]
RP   3D-STRUCTURE MODELING OF LECTIN DOMAIN.
RX   MEDLINE=94191542; PubMed=8142907;
RA   Padlan E.A., Helm B.A.;
RT   "Modeling of the lectin-homology domains of the human and murine low-
RT   affinity Fc epsilon receptor (Fc epsilon RII/CD23).";
RL   Receptor 3:325-341(1993).
RN   [7]
RP   3D-STRUCTURE MODELING OF 173-285.
RX   MEDLINE=96276216; PubMed=8745401;
RA   Bajorath J., Aruffo A.;
RT   "Structure-based modeling of the ligand binding domain of the human
RT   cell surface receptor CD23 and comparison of two independently
RT   derived molecular models.";
RL   Protein Sci. 5:240-247(1996).
CC   -!- FUNCTION: This receptor has essential roles in the regulation of
CC       IgE production and in the differentiation of B-cells (it is a B-
CC       cell-specific antigen).
CC   -!- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. ALSO EXISTS AS A
CC       SOLUBLE EXCRETED FORM.
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=2;
CC       Name=A;
CC         IsoId=P06734-1; Sequence=Displayed;
CC       Name=B;
CC         IsoId=P06734-2; Sequence=VSP_003057;
CC   -!- PTM: N- and O-glycosylated.
CC   -!- MISCELLANEOUS: There are two kinds of Fc receptors for IgE, which
CC       differ in both structure and function: high affinity receptors on
CC       basophils and mast cells and low affinity receptors on lymphocytes
CC       and monocytes.
CC   -!- SIMILARITY: Contains 1 C-type lectin family domain.
CC   -!- DATABASE: NAME=PROW; NOTE=CD guide CD23 entry;
CC       WWW="http://www.ncbi.nlm.nih.gov/prow/cd/cd23.htm".
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M15059; AAA52434.1; -.
DR   EMBL; M14766; AAA52435.1; -.
DR   EMBL; X04772; CAA28465.1; -.
DR   EMBL; M23562; AAA52433.1; -.
DR   PIR; A26067; LNHUER.
DR   PDB; 1HLI; 31-JAN-94.
DR   PDB; 1KJE; 03-APR-96.
DR   Genew; HGNC:3612; FCER2.
DR   MIM; 151445; -.
DR   GO; GO:0005887; C:integral to plasma membrane; TAS.
DR   GO; GO:0005178; F:integrin binding; TAS.
DR   InterPro; IPR002353; AntifreezeII.
DR   InterPro; IPR001304; Lectin_C.
DR   Pfam; PF00059; lectin_c; 1.
DR   PRINTS; PR00356; ANTIFREEZEII.
DR   SMART; SM00034; CLECT; 1.
DR   PROSITE; PS00615; C_TYPE_LECTIN_1; 1.
DR   PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
KW   Receptor; Antigen; IgE-binding protein; Repeat; Signal-anchor;
KW   Transmembrane; Lectin; Glycoprotein; Alternative splicing;
KW   3D-structure.
```

```
FT   CHAIN        1    321       MEMBRANE BOUND FORM.
FT   CHAIN      150    321       SOLUBLE FORM.
FT   DOMAIN       1     21       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    22     47       SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
FT                               (POTENTIAL).
FT   DOMAIN      48    321       EXTRACELLULAR (POTENTIAL).
FT   DOMAIN     162    284       C-TYPE LECTIN (LONG FORM).
FT   SITE       149    150       CLEAVAGE.
FT   REPEAT      69     89
FT   REPEAT      90    110
FT   REPEAT     111    131
FT   DISULFID   160    288       BY SIMILARITY.
FT   DISULFID   163    174       BY SIMILARITY.
FT   DISULFID   191    282
FT   DISULFID   259    273
FT   CARBOHYD    63     63       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   VARSPLIC     1      7       MEEGQYS -> MNPPSQ (in isoform B).
FT                               /FTId=VSP_003057.
FT   CONFLICT   269    269       N -> T (IN REF. 3).
FT   STRAND     174    177
FT   HELIX      184    193
FT   TURN       194    195
FT   STRAND     197    198
FT   HELIX      204    214
FT   TURN       215    216
FT   STRAND     219    228
FT   TURN       229    230
FT   STRAND     231    234
FT   TURN       235    236
FT   STRAND     239    239
FT   STRAND     245    245
FT   TURN       247    248
FT   TURN       254    255
FT   STRAND     259    262
FT   TURN       264    265
FT   STRAND     268    271
FT   TURN       273    274
FT   STRAND     281    284
SQ   SEQUENCE   321 AA;  36468 MW;  F86708C0E6515B87 CRC64;

  Query Match             13.4%;  Score 149;  DB 1;  Length 321;
  Best Local Similarity   30.2%;  Pred. No. 4.6e-06;
  Matches   55;  Conservative   20;  Mismatches   57;  Indels   50;  Gaps   11;

Qy         17 QRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFW 76
              || ||   ||   :::  :  |: ||    |||||| | :||  : |       |    |
Db        171 QRKCY---YFGKGTKQ--WVHARYACDDMEGQLVSIHSPEEQDFLTKH-----ASHTGSW 220

Qy         77 IGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEPSCGS--EVCVVMYHQPSAPAGI 134
              ||||  : |          : | |||   : ||   ||:  |  | ||:|       |
Db        221 IGLRNLDLKGE--------FIWVDGSHVDYSNWAPGEPTSRSQGEDCVMM-------RGS 265

Qy        135 GGPYMFQWNDDRCNMK-NNFICKYSDEKPAV---PSREAEGE----------ETELTTPV 180
              |     :|||  |: |   ::|    :: |    |: |   |          :  | ||
Db        266 G-----RWNDAFCDRKLGAWVC----DRLATCTPPASEGSAESMGPDSRPDPDGRLPTPS 316

Qy        181 LP 182
              ||
Db        317 AP 318


RESULT 23
PGCA_MOUSE
ID   PGCA_MOUSE     STANDARD;      PRT;  2132 AA.
AC   Q61282; Q64021;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   Aggrecan core protein precursor (Cartilage-specific proteoglycan core
DE   protein) (CSPCP).
GN   AGC1 OR AGC.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

```
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BALB/c; TISSUE=Cartilage;
RX   MEDLINE=95104847; PubMed=7806222;
RA   Walcz E., Deak F., Erhardt P., Coulter S.N., Fueloep C., Horvath P.,
RA   Doege K.J., Glant T.T.;
RT   "Complete coding sequence, deduced primary structure, chromosomal
RT   localization, and structural analysis of murine aggrecan.";
RL   Genomics 22:364-371(1994).
RN   [2]
RP   SEQUENCE OF 211-326 FROM N.A.
RC   STRAIN=129/Sv;
RX   MEDLINE=95004579; PubMed=7920633;
RA   Watanabe H., Kimata K., Line S., Strong D., Gao L.-Y., Kozak C.A.,
RA   Yamada Y.;
RT   "Mouse cartilage matrix deficiency (cmd) caused by a 7 bp deletion in
RT   the aggrecan gene.";
RL   Nat. Genet. 7:154-157(1994).
RN   [3]
RP   INTERACTION WITH FBLN1.
RX   MEDLINE=99329059; PubMed=10400671;
RA   Aspberg A., Adam S., Kostka G., Timpl R., Heinegaard D.;
RT   "Fibulin-1 is a ligand for the C-type lectin domains of aggrecan and
RT   versican.";
RL   J. Biol. Chem. 274:20444-20449(1999).
CC   -!- FUNCTION: This proteoglycan is a major component of extracellular
CC       matrix of cartilagenous tissues. A major function of this protein
CC       is to resist compression in cartilage. It binds avidly to
CC       hyaluronic acid via an amino-terminal globular region. May play a
CC       regulatory role in the matrix assembly of the cartilage.
CC   -!- SUBUNIT: Interacts with FBLN1.
CC   -!- SUBCELLULAR LOCATION: Secreted; extracellular matrix (By
CC       similarity).
CC   -!- DOMAIN: Two globular domains, G1 and G2, comprise the amino
CC       terminus of the proteoglycan, while another globular region, G3,
CC       makes up the COOH terminus. G1 contains link domains and thus
CC       consists of three disulfide-bonded loop structures designated as
CC       the A, B, B' motifs. G2 is similar to G1. The keratan sulfate (KS)
CC       and the chondroitin sulfate (CS) attachment domains lie between G2
CC       and G3.
CC   -!- PTM: Contains mostly chondroitin sulfate, but also keratan sulfate
CC       chains, N-linked and O-linked oligosaccharides.
CC   -!- DISEASE: Defects in AGC1 are the cause of cartilage matrix
CC       deficiency (CMD). CMD is an autosomal recessive syndrome
CC       characterized by cleft palate, short limbs, tail and snout.
CC       Mutation in strain CMD causes absence of aggrecan by truncation of
CC       the protein (mutation in the G1 domain).
CC   -!- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
CC   -!- SIMILARITY: Contains 4 link domains.
CC   -!- SIMILARITY: Contains 1 EGF-like domain.
CC   -!- SIMILARITY: Contains 1 C-type lectin family domain.
CC   -!- SIMILARITY: Contains 1 Sushi (SCR) domain.
CC   -!- SIMILARITY: Belongs to the aggrecan/versican proteoglycan family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; L07049; AAC37670.1; -.
DR   EMBL; S73722; AAB32160.1; -.
DR   EMBL; S73721; AAB32160.1; JOINED.
DR   PIR; A55182; A55182.
DR   HSSP; P98066; 1TSG.
DR   MGD; MGI:99602; Agc1.
DR   InterPro; IPR002353; AntifreezeII.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR003596; Ig_v.
DR   InterPro; IPR001304; Lectin_C.
DR   InterPro; IPR000538; Link.
DR   InterPro; IPR003324; SGXXSG.
DR   InterPro; IPR000436; Sushi_SCR_CCP.
DR   Pfam; PF00047; ig; 1.
DR   Pfam; PF00059; lectin_c; 1.
DR   Pfam; PF02339; SGXXSG; 60.
DR   Pfam; PF00084; sushi; 1.
DR   Pfam; PF00193; Xlink; 4.
DR   PRINTS; PR00356; ANTIFREEZEII.
DR   PRINTS; PR01265; LINKMODULE.
DR   ProDom; PD000918; Link; 4.
DR   SMART; SM00032; CCP; 1.
DR   SMART; SM00034; CLECT; 1.
DR   SMART; SM00406; IGv; 1.
DR   SMART; SM00445; LINK; 4.
DR   PROSITE; PS00615; C_TYPE_LECTIN_1; 1.
DR   PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
DR   PROSITE; PS50835; IG_LIKE; 1.
DR   PROSITE; PS00290; IG_MHC; 1.
DR   PROSITE; PS01241; LINK; 4.
KW   Glycoprotein; Proteoglycan; Lectin; Signal; Sushi; Repeat;
KW   Immunoglobulin domain.
FT   SIGNAL        1     19       POTENTIAL.
FT   CHAIN        20   2132       AGGRECAN CORE PROTEIN.
FT   DOMAIN       34    147       IG-LIKE V-TYPE.
FT   DOMAIN      170    247       LINK 1.
FT   DOMAIN      268    349       LINK 2.
FT   DOMAIN      504    581       LINK 3.
FT   DOMAIN      602    683       LINK 4.
FT   DOMAIN     1918   2044       C-TYPE LECTIN.
FT   DOMAIN     2048   2106       SUSHI.
FT   DOMAIN       48    140       G1-A.
FT   DOMAIN      152    247       G1-B.
FT   DOMAIN      253    349       G1-B'.
FT   DOMAIN      486    580       G2-B.
FT   DOMAIN      587    682       G2-B'.
FT   DOMAIN      685    803       KS.
FT   DOMAIN      805   1231       CS-1.
FT   DOMAIN     1232   1917       CS-2.
FT   DOMAIN     1917   2132       G3.
FT   DISULFID     51    133       BY SIMILARITY.
FT   DISULFID    175    246       BY SIMILARITY.
FT   DISULFID    199    220       BY SIMILARITY.
FT   DISULFID    273    348       BY SIMILARITY.
FT   DISULFID    297    318       BY SIMILARITY.
FT   DISULFID    509    580       BY SIMILARITY.
FT   DISULFID    533    554       BY SIMILARITY.
FT   DISULFID    607    682       BY SIMILARITY.
FT   DISULFID    631    652       BY SIMILARITY.
FT   DISULFID   1922   1933       BY SIMILARITY.
FT   DISULFID   1950   2042       BY SIMILARITY.
FT   DISULFID   2018   2034       BY SIMILARITY.
FT   DISULFID   2049   2092       BY SIMILARITY.
FT   DISULFID   2078   2105       BY SIMILARITY.
FT   CARBOHYD    126    126       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    239    239       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    333    333       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    387    387       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    611    611       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    667    667       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1675   1675       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   SITE       1171   1173       CELL ATTACHMENT SITE (POTENTIAL).
SQ   SEQUENCE   2132 AA;  222008 MW;  0B2BCDFC6CBDA163 CRC64;

  Query Match             13.3%;  Score 148.5;  DB 1;  Length 2132;
  Best Local Similarity   24.7%;  Pred. No. 5.3e-05;
  Matches   46;  Conservative   30;  Mismatches   67;  Indels   43;  Gaps    9;

Qy          2 TGRLLSGQPVCRGG---TQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQ 58
              |   :: |  |  |    |  ||:  :| |   |  : :|:  ||     | || : :||
Db       1912 TESTVADQEQCEEGWTKFQGHCYR--HFPD---RETWVDAERRCREQQSHLSSIVTPEEQ 1966
```

```
Qy          59 KLIEKFIENLLPSDGDF-WIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEP--- 114
               : : |  :       |: ||||         |    :  : |:||   ||  |  ::|
Db        1967 EFVNKNAQ-------DYQWIGL--------NDRTIEGDFRWSDGHSLQFEKWRPNQPDNF 2011

Qy         115 -SCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYS----DEKPAVPSREA 169
                : | :  |:::|:           :|||  || :  | ||       : | |
Db        2012 FATGEDCVVMIWHERG----------EWNDVPCNYQLPFTCKKGTVACGDPPVVEHART 2060

Qy         170 EGEETE 175
                |:: :
Db        2061 LGQKKD 2066


RESULT 24
PGCB_MOUSE
ID   PGCB_MOUSE     STANDARD;      PRT;   883 AA.
AC   Q61361;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Brevican core protein precursor.
GN   BCAN.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BALB/c; TISSUE=Brain;
RX   MEDLINE=97432816; PubMed=9286696;
RA   Rauch U., Meyer H., Brakebusch C., Seidenbecher C., Gundelfinger E.D.,
RA   Beier D.R., Fassler R.;
RT   "Sequence and chromosomal localization of the mouse brevican gene.";
RL   Genomics 44:15-21(1997).
CC   -!- FUNCTION: May play a role in the terminally differentiating and
CC       the adult nervous system during postnatal development. Could
CC       stabilize interactions between HA and brain proteoglycans.
CC   -!- SUBCELLULAR LOCATION: Secreted; extracellular matrix (By
CC       similarity).
CC   -!- TISSUE SPECIFICITY: Brain (By similarity).
CC   -!- PTM: Contains mostly chondroitin sulfate (By similarity).
CC   -!- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
CC   -!- SIMILARITY: Contains 2 link domains.
CC   -!- SIMILARITY: Contains 1 EGF-like domain.
CC   -!- SIMILARITY: Contains 1 C-type lectin family domain.
CC   -!- SIMILARITY: Contains 1 Sushi (SCR) domain.
CC   -!- SIMILARITY: Belongs to the aggrecan/versican proteoglycan family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X87096; CAA60575.1; -.
DR   PIR; S57653; S57653.
DR   HSSP; P20693; 1HLJ.
DR   MGD; MGI:1096385; Bcan.
DR   InterPro; IPR002353; AntifreezeII.
DR   InterPro; IPR000742; EGF_2.
DR   InterPro; IPR006209; EGF_like.
DR   InterPro; IPR006210; IEGF.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR003596; Ig_v.
DR   InterPro; IPR001304; Lectin_C.
DR   InterPro; IPR000538; Link.
DR   InterPro; IPR000436; Sushi_SCR_CCP.
DR   Pfam; PF00008; EGF; 1.
DR   Pfam; PF00047; ig; 1.
DR   Pfam; PF00059; lectin_c; 1.
DR   Pfam; PF00084; sushi; 1.
DR   Pfam; PF00193; Xlink; 2.
DR   PRINTS; PR00356; ANTIFREEZEII.
DR   PRINTS; PR01265; LINKMODULE.
DR   ProDom; PD000918; Link; 2.
DR   SMART; SM00032; CCP; 1.
DR   SMART; SM00034; CLECT; 1.
DR   SMART; SM00181; EGF; 1.
DR   SMART; SM00406; IGv; 1.
DR   SMART; SM00445; LINK; 2.
DR   PROSITE; PS00615; C_TYPE_LECTIN_1; 1.
DR   PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
DR   PROSITE; PS00022; EGF_1; 1.
DR   PROSITE; PS01186; EGF_2; 1.
DR   PROSITE; PS50026; EGF_3; 1.
DR   PROSITE; PS50835; IG_LIKE; 1.
DR   PROSITE; PS00290; IG_MHC; 1.
DR   PROSITE; PS01241; LINK; 2.
KW   Glycoprotein; Hyaluronic acid; Proteoglycan; Lectin; Signal; Sushi;
KW   EGF-like domain; Repeat; Immunoglobulin domain.
FT   SIGNAL        1     22       POTENTIAL.
FT   CHAIN        23    883       BREVICAN CORE PROTEIN.
FT   DOMAIN       35    154       IG-LIKE V-TYPE.
FT   DOMAIN      173    250       LINK 1.
FT   DOMAIN      271    352       LINK 2.
FT   DOMAIN      622    658       EGF-LIKE.
FT   DOMAIN      658    786       C-TYPE LECTIN.
FT   DOMAIN      787    851       SUSHI.
FT   DISULFID     56    136       BY SIMILARITY.
FT   DISULFID    178    249       BY SIMILARITY.
FT   DISULFID    202    223       BY SIMILARITY.
FT   DISULFID    276    351       BY SIMILARITY.
FT   DISULFID    300    321       BY SIMILARITY.
FT   DISULFID    626    637       BY SIMILARITY.
FT   DISULFID    631    646       BY SIMILARITY.
FT   DISULFID    648    657       BY SIMILARITY.
FT   DISULFID    664    675       BY SIMILARITY.
FT   DISULFID    692    784       BY SIMILARITY.
FT   DISULFID    760    776       BY SIMILARITY.
FT   DISULFID    791    834       BY SIMILARITY.
FT   DISULFID    820    847       BY SIMILARITY.
FT   CARBOHYD    129    129       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    336    336       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   883 AA;  96013 MW;  CC2C33C97B453E45 CRC64;

  Query Match             13.2%;  Score 147;  DB 1;  Length 883;
  Best Local Similarity   29.7%;  Pred. No. 2.4e-05;
  Matches   43;  Conservative   21;  Mismatches   45;  Indels   36;  Gaps    8;

Qy          17 QRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLI-EKFIENLLPSDGDF 75
               |  |||    | ::|| ::|||:  ||  |  | || : :||  : ::: |
Db         672 QGACYK----HFSTRR-SWEEAESQCRALGAHLTSICTPEEQDFVNDRYREYQ------- 719

Qy          76 WIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEPS---CGSEVCVVM-YHQPSAP 131
               ||||        |    :  : |:||:   : ||   :|       | |||| :|
Db         720 WIGL--------NDRTIEGDFLWSDGAPLLYENWNPGQPDSYFLSGENCVVMVWHDQG-- 769

Qy         132 AGIGGPYMFQWNDDRCNMKNNFICK 156
                       ||:|  ||   :: ||
Db         770 ---------QWSDVPCNYHLSYTCK 785


RESULT 25
PGCA_CHICK
ID   PGCA_CHICK     STANDARD;      PRT;  2109 AA.
AC   P07898; Q90810; Q90820; Q90991; Q91047;
DT   01-AUG-1988 (Rel. 08, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   Aggrecan core protein precursor (Cartilage-specific proteoglycan core
DE   protein) (CSPCP).
GN   AGC1.
```

```
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
OX   NCBI_TaxID=9031;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=White leghorn; TISSUE=Embryo;
RX   MEDLINE=94043149; PubMed=8226878;
RA   Li H., Schwartz N.B., Vertel B.M.;
RT   "cDNA cloning of chick cartilage chondroitin sulfate (aggrecan) core
RT   protein and identification of a stop codon in the aggrecan gene
RT   associated with the chondrodystrophy, nanomelia.";
RL   J. Biol. Chem. 268:23504-23511(1993).
RN   [2]
RP   SEQUENCE OF 1042-1559 FROM N.A.
RC   TISSUE=Embryo;
RX   MEDLINE=90307744; PubMed=1694853;
RA   Krueger R.C. Jr., Fields T.A., Mensch J.R. Jr., Schwartz N.B.;
RT   "Chick cartilage chondroitin sulfate proteoglycan core protein. II.
RT   Nucleotide sequence of cDNA clone and localization of the S103L
RT   epitope.";
RL   J. Biol. Chem. 265:12088-12097(1990).
RN   [3]
RP   SEQUENCE OF 1-1855 AND 1893-2109 FROM N.A.
RC   TISSUE=Cartilage;
RX   MEDLINE=93111968; PubMed=1339285;
RA   Chandrasekaran L., Tanzer M.L.;
RT   "Molecular cloning of chicken aggrecan. Structural analyses.";
RL   Biochem. J. 288:903-910(1992).
RN   [4]
RP   ERRATUM.
RX   MEDLINE=94107258; PubMed=8280087;
RA   Chandrasekaran L., Tanzer M.L.;
RL   Biochem. J. 296:885-887(1993).
RN   [5]
RP   SEQUENCE OF 1492-1610 FROM N.A.
RC   STRAIN=White leghorn; TISSUE=Chondrocytes;
RX   MEDLINE=95128519; PubMed=7827752;
RA   Primorac D., Stover M.L., Clark S.H., Rowe D.W.;
RT   "Molecular basis of nanomelia, a heritable chondrodystrophy of
RT   chicken.";
RL   Matrix Biol. 14:297-305(1994).
RN   [6]
RP   SEQUENCE OF 1894-2109 FROM N.A.
RX   MEDLINE=89008500; PubMed=3170613;
RA   Tanaka T., Har-El R., Tanzer M.L.;
RT   "Partial structure of the gene for chicken cartilage proteoglycan
RT   core protein.";
RL   J. Biol. Chem. 263:15831-15835(1988).
RN   [7]
RP   SEQUENCE OF 1693-1855 AND 1893-2109 FROM N.A.
RX   MEDLINE=86259736; PubMed=3460082;
RA   Sai S., Tanaka T., Kosher R.A., Tanzer M.L.;
RT   "Cloning and sequence analysis of a partial cDNA for chicken
RT   cartilage proteoglycan core protein.";
RL   Proc. Natl. Acad. Sci. U.S.A. 83:5081-5085(1986).
CC   -!- FUNCTION: This proteoglycan is a major component of extracellular
CC       matrix of cartilagenous tissues. A major function of this protein
CC       is to resist compression in cartilage. It binds avidly to
CC       hyaluronic acid via an amino-terminal globular region. May play a
CC       regulatory role in the matrix assembly of the cartilage.
CC   -!- SUBCELLULAR LOCATION: Secreted; extracellular matrix (By
CC       similarity).
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=2;
CC       Name=1;
CC         IsoId=P07898-1; Sequence=Displayed;
CC       Name=2;
CC         IsoId=P07898-2; Sequence=VSP_003073;
CC   -!- DOMAIN: Two globular domains, G1 and G2, comprise the amino
CC       terminus of the proteoglycan, while another globular region, G3,
CC       makes up the COOH terminus. G1 contains link domains and thus
CC       consists of three disulfide-bonded loop structures designated as
CC       the A, B, B' motifs. G2 is similar to G1. The keratan sulfate (KS)
CC       and the chondroitin sulfate (CS) attachment domains lie between G2
CC       and G3.
CC   -!- PTM: Contains mostly chondroitin sulfate, but also keratan sulfate
CC       chains, N-linked and O-linked oligosaccharides.
CC   -!- DISEASE: DEFECTS IN AGC1 ARE THE CAUSE OF NANOMELIA, A LETHAL
CC       CONNECTIVE TISSUE DISORDER AFFECTING CARTILAGE DEVELOPMENT
CC       (CHONDRODYSTROPHY) CHARACTERIZED BY SHORTENED AND MALFORMED LIMBS.
CC       AGGRECAN IS TRUNCATED AT ITS C-TERMINAL IN THE CS-2 BINDING DOMAIN
CC       AND IS NOT ANYMORE SECRETED FROM THE CHONDROCYTES.
CC   -!- SIMILARITY: Belongs to the aggrecan/versican proteoglycan family.
CC   -!- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
CC   -!- SIMILARITY: Contains 4 link domains.
CC   -!- SIMILARITY: Contains 1 EGF-like domain.
CC   -!- SIMILARITY: Contains 1 C-type lectin family domain.
CC   -!- SIMILARITY: Contains 1 Sushi (SCR) domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; L21913; AAB19128.1; -.
DR   EMBL; M38187; AAA48731.1; -.
DR   EMBL; M88101; -; NOT_ANNOTATED_CDS.
DR   EMBL; S74657; AAC60751.1; -.
DR   EMBL; S74656; AAC60751.1; JOINED.
DR   EMBL; J04028; AAA48719.1; -.
DR   EMBL; M13993; AAA48720.1; -.
DR   PIR; I50421; I50421.
DR   HSSP; P08709; 1BF9.
DR   InterPro; IPR002353; AntifreezeII.
DR   InterPro; IPR000152; Asx_hydroxyl_S.
DR   InterPro; IPR000742; EGF_2.
DR   InterPro; IPR001881; EGF_Ca.
DR   InterPro; IPR006209; EGF_like.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003599; Ig.
DR   InterPro; IPR001304; Lectin_C.
DR   InterPro; IPR000538; Link.
DR   InterPro; IPR003324; SGXXSG.
DR   InterPro; IPR000436; Sushi_SCR_CCP.
DR   Pfam; PF00008; EGF; 1.
DR   Pfam; PF00047; ig; 1.
DR   Pfam; PF00059; lectin_c; 1.
DR   Pfam; PF02339; SGXXSG; 56.
DR   Pfam; PF00084; sushi; 1.
DR   Pfam; PF00193; Xlink; 4.
DR   PRINTS; PR00356; ANTIFREEZEII.
DR   PRINTS; PR01265; LINKMODULE.
DR   ProDom; PD000918; Link; 4.
DR   SMART; SM00032; CCP; 1.
DR   SMART; SM00034; CLECT; 1.
DR   SMART; SM00179; EGF_CA; 1.
DR   SMART; SM00409; IG; 1.
DR   SMART; SM00445; LINK; 4.
DR   PROSITE; PS00010; ASX_HYDROXYL; 1.
DR   PROSITE; PS00615; C_TYPE_LECTIN_1; 1.
DR   PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
DR   PROSITE; PS00022; EGF_1; 1.
DR   PROSITE; PS50026; EGF_3; 1.
DR   PROSITE; PS01187; EGF_CA; 1.
DR   PROSITE; PS50835; IG_LIKE; 1.
DR   PROSITE; PS01241; LINK; 4.
KW   Glycoprotein; Proteoglycan; Lectin; Signal; Sushi; EGF-like domain;
KW   Alternative splicing; Repeat; Immunoglobulin domain.
FT   SIGNAL        1     16       POTENTIAL.
FT   CHAIN        17   2109       AGGRECAN CORE PROTEIN.
FT   DOMAIN       34    143       IG-LIKE V-TYPE.
```

```
FT   DOMAIN      166    243       LINK 1.
FT   DOMAIN      264    346       LINK 2.
FT   DOMAIN      537    614       LINK 3.
FT   DOMAIN      635    716       LINK 4.
FT   DOMAIN     1363   1742       19 X 20 AA TANDEM-REPEAT.
FT   DOMAIN     1855   1892       EGF-LIKE.
FT   DOMAIN     1901   2019       C-TYPE LECTIN.
FT   DOMAIN     2023   2081       SUSHI.
FT   DOMAIN       48    137       G1-A.
FT   DOMAIN      148    243       G1-B.
FT   DOMAIN      249    346       G1-B'.
FT   DOMAIN      519    613       G2-B.
FT   DOMAIN      620    715       G2-B'.
FT   DOMAIN      718    803       KS.
FT   DOMAIN      805   1264       CS-1.
FT   DOMAIN     1265   1742       CS-2.
FT   DOMAIN     1893   2109       G3.
FT   DISULFID     51    129       BY SIMILARITY.
FT   DISULFID    171    242       BY SIMILARITY.
FT   DISULFID    195    216       BY SIMILARITY.
FT   DISULFID    269    345       BY SIMILARITY.
FT   DISULFID    293    314       BY SIMILARITY.
FT   DISULFID    542    613       BY SIMILARITY.
FT   DISULFID    566    587       BY SIMILARITY.
FT   DISULFID    640    715       BY SIMILARITY.
FT   DISULFID    664    685       BY SIMILARITY.
FT   DISULFID   1859   1870       BY SIMILARITY.
FT   DISULFID   1864   1879       BY SIMILARITY.
FT   DISULFID   1881   1890       BY SIMILARITY.
FT   DISULFID   1897   1908       BY SIMILARITY.
FT   DISULFID   1925   2017       BY SIMILARITY.
FT   DISULFID   1993   2009       BY SIMILARITY.
FT   DISULFID   2024   2067       BY SIMILARITY.
FT   DISULFID   2053   2080       BY SIMILARITY.
FT   CARBOHYD     76     76       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    122    122       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    330    330       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    388    388       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    439    439       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    644    644       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    700    700       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    765    765       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    801    801       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   VARSPLIC   1856   1892       Missing (in isoform 2).
FT                                /FTId=VSP_003073.
FT   CONFLICT    362    362       E -> D (IN REF. 3).

  Query Match             13.1%;  Score 146;  DB 1;  Length 2109;
  Best Local Similarity   25.0%;  Pred. No. 8.7e-05;
  Matches   42;  Conservative   28;  Mismatches   58;  Indels   40;  Gaps    8;

Qy         17 QRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDF- 75
              |  ||:  :| :   |  : :|:  ||      | || : :||: :    :       |:
Db       1905 QGHCYR--HFEE---RETWMDAESRCREHQAHLSSIITPEEQEFVNSHAQ-------DYQ 1952

Qy         76 WIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEPS----CGSEVCVVMYHQPSAP 131
              ||||  |        | :: : |:||   || ||  ::|      | :  |:::|:
Db       1953 WIGLSDR--------AVENDFRWSDGHSLQFENWRPNQPDNFFFAGEDCVVMIWHEQG-- 2002

Qy        132 AGIGGPYMFQWNDDRCNMKNNFICKYS----DEKPAVPSREAEGEETE 175
                  :|||  ||    | ||      : | | :    | : :
Db       2003 ---------EWNDVPCNYHLPFTCKKGTVACGDPPVVENARTFGRKKD 2041


RESULT 26
LECG_TRIST
ID   LECG_TRIST     STANDARD;      PRT;   158 AA.
AC   Q9YGP1;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Galactose-binding lectin precursor (TSL).
OS   Trimeresurus stejnegeri (Chinese green tree viper).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
OC   Viperidae; Crotalinae; Trimeresurus.
OX   NCBI_TaxID=39682;
RN   [1]
RP   SEQUENCE FROM N.A., SEQUENCE OF 24-53, AND MASS SPECTROMETRY.
RC   TISSUE=Venom;
RX   MEDLINE=99348038; PubMed=10417338;
RA   Xu Q., Wu X.-F., Xia Q.-C., Wang K.-Y.;
RT   "Cloning of a galactose-binding lectin from the venom of Trimeresurus
RT   stejnegeri.";
RL   Biochem. J. 341:733-737(1999).
RN   [2]
RP   SEQUENCE, CARBOHYDRATE-LINKAGE SITE, AND STRUCTURE OF CARBOHYDRATES.
RC   TISSUE=Venom;
RX   PubMed=10561575;
RA   Zeng R., Xu Q., Shao X.-X., Wang K.-Y., Xia Q.-C.;
RT   "Characterization and analysis of a novel glycoprotein from snake
RT   venom using liquid chromatography-electrospray mass spectrometry and
RT   Edman degradation.";
RL   Eur. J. Biochem. 266:352-358(1999).
CC   -!- FUNCTION: Galactose-binding protein which recognizes specific
CC       carbohydrate structures and agglutinate a variety of animal cells
CC       by binding to cell-surface glycoproteins and glycolipids. May be a
CC       calcium-dependent lectin.
CC   -!- SUBUNIT: Homodimer; disulfide-linked.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- MASS SPECTROMETRY: MW=17924.2; MW_ERR=2.4; METHOD=Electrospray;
CC       RANGE=24-158.
CC   -!- MISCELLANEOUS: Met-33 has been shown to be oxidized to methionine
CC       sulfoxide (Ref.2) but this probably results from sample treatment
CC       prior to mass spectrometry.
CC   -!- SIMILARITY: Contains 1 C-type lectin family domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF119097; AAD17252.1; -.
DR   HSSP; P22897; 1EGG.
DR   GlycoSuiteDB; Q9YGP1; -.
DR   InterPro; IPR001304; Lectin_C.
DR   InterPro; IPR003990; Pancreatis_ac.
DR   Pfam; PF00059; lectin_c; 1.
DR   PRINTS; PR01504; PNCREATITSAP.
DR   SMART; SM00034; CLECT; 1.
DR   PROSITE; PS00615; C_TYPE_LECTIN_1; 1.
DR   PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
KW   Lectin; Signal; Calcium; Glycoprotein.
FT   SIGNAL        1     23
FT   CHAIN        24    158       GALACTOSE-BINDING LECTIN.
FT   DOMAIN       24    158       C-TYPE LECTIN.
FT   DISULFID     26     37       BY SIMILARITY.
FT   DISULFID     54    154       BY SIMILARITY.
FT   DISULFID     61    156       BY SIMILARITY.
FT   DISULFID    129    146       BY SIMILARITY.
FT   DISULFID    109    109       INTERCHAIN (BY SIMILARITY).
FT   CARBOHYD     28     28       N-LINKED (GLCNAC. . .) (HIGH MANNOSE).
FT                                /FTId=CAR_000165.
SQ   SEQUENCE   158 AA;  18635 MW;  2077BC62B7A08FF9 CRC64;

  Query Match             13.0%;  Score 145;  DB 1;  Length 158;
  Best Local Similarity   26.7%;  Pred. No. 4.4e-06;
  Matches   39;  Conservative   20;  Mismatches   51;  Indels   36;  Gaps    5;

Qy         20 CYKVIYFHDTSRRLNFEEAKEACR--RDGGQLVSIESEDEQKLIEKFIENLLPSDGDFWI 77
              |||:          :|:|:  ||  : |  | |     |   | ::| :       : ||
Db         37 CYKIF-----DEPKTWEDAEMFCRKYKPGCHLASFHRLAESLDIAEYISDYHKRQAEVWI 91
```

```
Qy          78 GLRRREEKQSNSTACQDLYAWTDGSISQFRNW-------YVDEPSCGSEVCVVMYHQPSA 130
               ||  |::  |        : ||| | : : ||       | |:  |   | :  ||
Db          92 GLLDRKKDFS--------WEWTDRSCTDYLNWDKNQPDHYKDKEFCVELVSLTGYH---- 139

Qy         131 PAGIGGPYMFQWNDDRCNMKNNFICK 156
                     :|||  |  ||:|:|:
Db         140 ----------RWNDQVCESKNSFLCQ 155


RESULT 27
PGCB_RAT
ID   PGCB_RAT       STANDARD;      PRT;   883 AA.
AC   P55068; Q62860; Q63040; Q63513;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Brevican core protein precursor (Brain enriched hyaluronan binding
DE   protein) (BEHAB protein).
GN   BCAN.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Sprague-Dawley; TISSUE=Brain;
RX   MEDLINE=96070828; PubMed=7592978;
RA   Seidenbecher C.I., Richter K., Rauch U., Faessler R., Garner C.C.,
RA   Gundelfinger E.D.;
RT   "Brevican, a chondroitin sulfate proteoglycan of rat brain, occurs as
RT   secreted and cell surface glycosylphosphatidylinositol-anchored
RT   isoforms.";
RL   J. Biol. Chem. 270:27206-27212(1995).
RN   [2]
RP   SEQUENCE FROM N.A., AND SEQUENCE OF 396-407.
RC   TISSUE=Brain;
RX   MEDLINE=96074575; PubMed=7488217;
RA   Yamada H., Watanabe K., Shimonaka M., Yamasaki M., Yamaguchi Y.;
RT   "cDNA cloning and the identification of an aggrecanase-like cleavage
RT   site in rat brevican.";
RL   Biochem. Biophys. Res. Commun. 216:957-963(1995).
RN   [3]
RP   SEQUENCE OF 1-423 FROM N.A.
RC   STRAIN=Sprague-Dawley; TISSUE=Brain;
RX   MEDLINE=94216386; PubMed=7512973;
RA   Jaworski D.M., Kelly G.M., Hockfield S.;
RT   "BEHAB, a new member of the proteoglycan tandem repeat family of
RT   hyaluronan-binding proteins that is restricted to the brain.";
RL   J. Cell Biol. 125:495-509(1994).
CC   -!- FUNCTION: May play a role in the terminally differentiating and
CC       the adult nervous system during postnatal development. Could
CC       stabilize interactions between HA and brain proteoglycans. Isoform
CC       2 may function as a chondroitin sulfate-bearing cell surface
CC       receptor.
CC   -!- SUBCELLULAR LOCATION: Secreted (isoform 1); extracellular matrix
CC       and attached to the membrane by a GPI-anchor (isoform 2).
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=2;
CC       Name=1;
CC         IsoId=P55068-1; Sequence=Displayed;
CC       Name=2;
CC         IsoId=P55068-2; Sequence=VSP_003076, VSP_003077;
CC   -!- TISSUE SPECIFICITY: Brain.
CC   -!- DEVELOPMENTAL STAGE: Isoform 1 increases from day P4 to P64.
CC       Isoform 2 increases after day P8.
CC   -!- PTM: Contains mostly chondroitin sulfate.
CC   -!- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
CC   -!- SIMILARITY: Contains 2 link domains.
CC   -!- SIMILARITY: Contains 1 EGF-like domain.
CC   -!- SIMILARITY: Contains 1 C-type lectin family domain.
CC   -!- SIMILARITY: Contains 1 Sushi (SCR) domain.
CC   -!- SIMILARITY: Belongs to the aggrecan/versican proteoglycan family.
CC   -!- CAUTION: REF.3 SEQUENCE DIFFERS FROM THAT SHOWN FROM POSITION 364
CC       ONWARD AND IS SMALLER (371 AA) DUE TO A FRAMESHIFT.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X79881; CAA56255.1; -.
DR   EMBL; X86406; CAA60160.1; -.
DR   EMBL; U37142; AAA87847.1; -.
DR   EMBL; Z28366; CAA82215.1; ALT_FRAME.
DR   PIR; S49126; S49126.
DR   HSSP; P20693; 1HLJ.
DR   InterPro; IPR002353; AntifreezeII.
DR   InterPro; IPR000742; EGF_2.
DR   InterPro; IPR006209; EGF_like.
DR   InterPro; IPR006210; IEGF.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR003596; Ig_v.
DR   InterPro; IPR001304; Lectin_C.
DR   InterPro; IPR000538; Link.
DR   InterPro; IPR000436; Sushi_SCR_CCP.
DR   Pfam; PF00008; EGF; 1.
DR   Pfam; PF00047; ig; 1.
DR   Pfam; PF00059; lectin_c; 1.
DR   Pfam; PF00084; sushi; 1.
DR   Pfam; PF00193; Xlink; 2.
DR   PRINTS; PR00356; ANTIFREEZEII.
DR   PRINTS; PR01265; LINKMODULE.
DR   ProDom; PD000918; Link; 2.
DR   SMART; SM00032; CCP; 1.
DR   SMART; SM00034; CLECT; 1.
DR   SMART; SM00181; EGF; 1.
DR   SMART; SM00406; IGv; 1.
DR   SMART; SM00445; LINK; 2.
DR   PROSITE; PS00615; C_TYPE_LECTIN_1; 1.
DR   PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
DR   PROSITE; PS00022; EGF_1; 1.
DR   PROSITE; PS01186; EGF_2; 1.
DR   PROSITE; PS50026; EGF_3; 1.
DR   PROSITE; PS50835; IG_LIKE; 1.
DR   PROSITE; PS00290; IG_MHC; 1.
DR   PROSITE; PS01241; LINK; 2.
KW   Glycoprotein; Hyaluronic acid; Proteoglycan; Lectin; Signal; Sushi;
KW   EGF-like domain; Repeat; Immunoglobulin domain; Alternative splicing;
KW   GPI-anchor.
FT   SIGNAL        1     22       POTENTIAL.
FT   CHAIN        23    883       BREVICAN CORE PROTEIN.
FT   DOMAIN       35    154       IG-LIKE V-TYPE.
FT   DOMAIN      173    250       LINK 1.
FT   DOMAIN      271    352       LINK 2.
FT   DOMAIN      622    658       EGF-LIKE.
FT   DOMAIN      658    786       C-TYPE LECTIN.
FT   DOMAIN      787    851       SUSHI.
FT   DISULFID     56    136       BY SIMILARITY.
FT   DISULFID    178    249       BY SIMILARITY.
FT   DISULFID    202    223       BY SIMILARITY.
FT   DISULFID    276    351       BY SIMILARITY.
FT   DISULFID    300    321       BY SIMILARITY.
FT   DISULFID    626    637       BY SIMILARITY.
FT   DISULFID    631    646       BY SIMILARITY.
FT   DISULFID    648    657       BY SIMILARITY.
FT   DISULFID    791    834       BY SIMILARITY.
FT   DISULFID    820    847       BY SIMILARITY.
FT   CARBOHYD    129    129       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    336    336       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   VARSPLIC    625    645       DCIPSPCHNGGTCLEEKEGFR -> NSAEGSMPAFLLFLLL
FT                                QLWDT (in isoform 2).
FT                                /FTId=VSP_003076.
```

```
FT   VARSPLIC    646     883       Missing (in isoform 2).
FT                                 /FTId=VSP_003077.
FT   CONFLICT     51      52       AL -> WV (IN REF. 3).
FT   CONFLICT    503     503       V -> L (IN REF. 2).
FT   CONFLICT    518     519       TV -> PA (IN REF. 2).
FT   CONFLICT    526     526       G -> R (IN REF. 2).
FT   CONFLICT    541     541       G -> A (IN REF. 2).
FT   CONFLICT    556     556       R -> S (IN REF. 2).
FT   CONFLICT    573     573       E -> A (IN REF. 2).
FT   CONFLICT    583     583       V -> L (IN REF. 2).
FT   CONFLICT    649     649       V -> L (IN REF. 2).
FT   CONFLICT    670     670       P -> A (IN REF. 2).
FT   CONFLICT    738     738       P -> A (IN REF. 2).
FT   CONFLICT    809     809       R -> A (IN REF. 2).
SQ   SEQUENCE   883 AA;  96057 MW;  AC7ACC40CB53ED37 CRC64;
```

Query Match 13.0%; Score 145; DB 1; Length 883;
Best Local Similarity 29.7%; Pred. No. 3.7e-05;
Matches 43; Conservative 20; Mismatches 46; Indels 36; Gaps 8;

```
Qy         17 QRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLI-EKFIENLLPSDGDF 75
              |  |||    | ::|| ::|||:  ||  |  | || : :||  : ::: |
Db        672 QGACYK----HFSTRR-SWEEAESQCRALGAHLTSICTPEEQDFVNDRYREYQ------- 719

Qy         76 WIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEPS---CGSEVCVVM-YHQPSAP 131
              ||||        |    :  : |:||    : ||   :|       | |||| :|
Db        720 WIGL--------NDRTIEGDFLWSDGPPLLYENWNPGQPDSYFLSGENCVVMVWHDQG-- 769

Qy        132 AGIGGPYMFQWNDDRCNMKNNFICK 156
                    ||:|  ||   :: ||
Db        770 ---------QWSDVPCNYHLSYTCK 785
```

```
RESULT 28
LEC2_MEGRO
ID   LEC2_MEGRO     STANDARD;      PRT;   173 AA.
AC   P17346;
DT   01-AUG-1990 (Rel. 15, Created)
DT   01-AUG-1990 (Rel. 15, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Lectin BRA-2.
OS   Megabalanus rosa (Acorn barnacle).
OC   Eukaryota; Metazoa; Arthropoda; Crustacea; Maxillopoda; Cirripedia;
OC   Thoracica; Sessilia; Balanomorpha; Balanoidea; Balanidae;
OC   Megabalanus.
OX   NCBI_TaxID=6680;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Coelomic fluid;
RX   MEDLINE=90283457; PubMed=2354200;
RA   Muramoto K., Kamiya H.;
RT   "The amino-acid sequence of multiple lectins of the acorn barnacle
RT   Megabalanus rosa and its homology with animal lectins.";
RL   Biochim. Biophys. Acta 1039:42-51(1990).
RN   [2]
RP   DISULFIDE BONDS, AND CARBOHYDRATE-LINKAGE SITE.
RX   MEDLINE=90283459; PubMed=2354201;
RA   Muramoto K., Kamiya H.;
RT   "The positions of the disulfide bonds and the glycosylation site in a
RT   lectin of the acorn barnacle Megabalanus rosa.";
RL   Biochim. Biophys. Acta 1039:52-60(1990).
CC   -!- FUNCTION: Sugar-binding protein which recognizes specific
CC       carbohydrate structures and agglutinate a variety of animal cells
CC       by binding to cell-surface glycoproteins and glycolipids. Calcium-
CC       dependent lectin. Invertebrate lectins may be involved in defense
CC       functions.
CC   -!- SUBUNIT: Homohexamer; disulfide-linked.
CC   -!- TISSUE SPECIFICITY: Coelemic fluid.
CC   -!- MISCELLANEOUS: This lectin binds galactose.
CC   -!- SIMILARITY: Contains 1 C-type lectin family domain.
DR   PIR; S10548; S10548.
DR   HSSP; P23806; 1IXX.
DR   InterPro; IPR002353; AntifreezeII.
DR   InterPro; IPR001304; Lectin_C.
DR   Pfam; PF00059; lectin_c; 1.
DR   PRINTS; PR00356; ANTIFREEZEII.
DR   SMART; SM00034; CLECT; 1.
DR   PROSITE; PS00615; C_TYPE_LECTIN_1; 1.
DR   PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
KW   Glycoprotein; Lectin; Calcium.
FT   DOMAIN       51    170       C-TYPE LECTIN (LONG FORM).
FT   DISULFID     14     14       INTERCHAIN (WITH C-50 IN OTHER CHAIN).
FT   DISULFID     50     50       INTERCHAIN (WITH C-1 IN OTHER CHAIN).
FT   DISULFID     47     61
FT   DISULFID     78    168
FT   DISULFID    144    160
FT   CARBOHYD     39     39       N-LINKED (GLCNAC. . .).
SQ   SEQUENCE   173 AA;  19580 MW;  0D90DCA1F805D818 CRC64;
```

Query Match 13.0%; Score 144.5; DB 1; Length 173;
Best Local Similarity 28.6%; Pred. No. 5.4e-06;
Matches 38; Conservative 18; Mismatches 50; Indels 27; Gaps 5;

```
Qy         26 FHDTSRRLNFEEAKEACRR--DGGQLVSIESEDEQKLIEKFIENLLPSDGDFWIGLRRRE 83
              ||    : ::  |   | |     :| ||:: | | ::|        |    ||||
Db         62 FHVPLEKASWMVAHGVCARLDSRARLASIDAAD-QAVVEPL------SSEKMWIGL---- 110

Qy         84 EKQSNSTACQDLYAWTDGSISQFRNWYVDEPSCGSEVCVVMYHQPSAPAGIGGPYMFQWN 143
                 :  :  |   | | | |  ||||  :|   ||:||::                ||:
Db        111 ----SYDSANDAAVWADDSHSSHRNWYATQPDDESELCVLIKEDQYR----------QWH 156

Qy        144 DDRCNMKNNFICK 156
              |  || : ||:|:
Db        157 DYNCNDRYNFVCE 169
```

```
RESULT 29
LEM1_RAT
ID   LEM1_RAT       STANDARD;      PRT;   372 AA.
AC   P30836;
DT   01-JUL-1993 (Rel. 26, Created)
DT   01-JUL-1993 (Rel. 26, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   L-selectin precursor (Lymph node homing receptor) (Leukocyte adhesion
DE   molecule-1) (LAM-1) (LY-22) (Lymphocyte surface MEL-14 antigen)
DE   (Leukocyte-endothelial cell adhesion molecule 1) (LECAM1) (CD62L).
GN   SELL OR LNHR OR LY-22.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=92329548; PubMed=1378303;
RA   Watanabe T., Song Y., Hirayama Y., Tamatani T., Kuida K., Miyasaka M.;
RT   "Sequence and expression of a rat cDNA for LECAM-1.";
RL   Biochim. Biophys. Acta 1131:321-324(1992).
CC   -!- FUNCTION: Cell surface adhesion protein. Mediate the adherence of
CC       lymphocytes to endothelial cells of high endothelial venules in
CC       peripheral lymph nodes.
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC   -!- SIMILARITY: Belongs to the selectin/LECAM family.
CC   -!- SIMILARITY: Contains 1 C-type lectin family domain.
CC   -!- SIMILARITY: Contains 1 EGF-like domain.
CC   -!- SIMILARITY: Contains 2 Sushi (SCR) domains.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; D10831; BAA01613.1; -.
DR   PIR; S23936; S23936.
```

```
DR   HSSP; P14151; 1KJB.
DR   InterPro; IPR006209; EGF_like.
DR   InterPro; IPR006210; IEGF.
DR   InterPro; IPR001304; Lectin_C.
DR   InterPro; IPR002396; Selectin.
DR   InterPro; IPR000436; Sushi_SCR_CCP.
DR   Pfam; PF00008; EGF; 1.
DR   Pfam; PF00059; lectin_c; 1.
DR   Pfam; PF00084; sushi; 2.
DR   PRINTS; PR00343; SELECTIN.
DR   SMART; SM00032; CCP; 2.
DR   SMART; SM00034; CLECT; 1.
DR   SMART; SM00181; EGF; 1.
DR   PROSITE; PS00615; C_TYPE_LECTIN_1; 1.
DR   PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
DR   PROSITE; PS00022; EGF_1; 1.
DR   PROSITE; PS01186; EGF_2; 1.
DR   PROSITE; PS50026; EGF_3; 1.
KW   Cell adhesion; Transmembrane; Glycoprotein; EGF-like domain; Lectin;
KW   Selectin; Signal; Sushi; Repeat.
FT   SIGNAL       1     28       BY SIMILARITY.
FT   PROPEP      29     38       BY SIMILARITY.
FT   CHAIN       39    372       L-SELECTIN.
FT   DOMAIN      39    332       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   333    355       POTENTIAL.
FT   DOMAIN     356    372       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN      55    155       C-TYPE LECTIN (SHORT FORM).
FT   DOMAIN     156    192       EGF-LIKE.
FT   DOMAIN     196    255       SUSHI 1.
FT   DOMAIN     258    317       SUSHI 2.
FT   DISULFID    57    155       BY SIMILARITY.
FT   DISULFID   128    147       BY SIMILARITY.
FT   DISULFID   160    171       BY SIMILARITY.
FT   DISULFID   165    180       BY SIMILARITY.
FT   DISULFID   182    191       BY SIMILARITY.
FT   DISULFID   197    241       BY SIMILARITY.
FT   DISULFID   227    254       BY SIMILARITY.
FT   DISULFID   259    303       BY SIMILARITY.
FT   DISULFID   289    316       BY SIMILARITY.
FT   CARBOHYD    60     60       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   104    104       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   177    177       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   226    226       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   246    246       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   278    278       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   372 AA;  42441 MW;  3B88AE0F1E4D191A CRC64;

  Query Match             13.0%;  Score 144.5;  DB 1;  Length 372;
  Best Local Similarity   26.5%;  Pred. No. 1.4e-05;
  Matches   41;  Conservative   35;  Mismatches   52;  Indels   27;  Gaps    8;

Qy         26 FHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGD-FWIGLRRREE 84
              :| : | :|:| |:: |: :   ||:|::: |   || ::|  || :   :|||:|:
Db         41 YHYSERSMNWENARKFCKHNYTDLVAIQNKRE---IE-YLEKTLPKNPTYYWIGIRK--- 93

Qy         85 KQSNSTACQDLYAW--TDGSIS-QFRNWYVDEPS--CGSEVCVVMYHQPSAPAGIGGPYM 139
                 :  |  |: ::: :  ||   ||:     | || :|:    :|
Db         94 -------IGKTWTWVGTNKTLTKEAENWGTGEPNNKKSKEDCVEIYIKRERDSG------ 140

Qy        140 FQWNDDRCNMKNNFICKYSDEKPAVPSREAEGEET 174
              :|||| |: :   :|  :  :|   :|  |  ||
Db        141 -KWNDDACHKRKAALCYTASCQPESCNRHGECVET 174


RESULT 30
LEM1_MOUSE
ID   LEM1_MOUSE     STANDARD;      PRT;   372 AA.
AC   P18337;
DT   01-NOV-1990 (Rel. 16, Created)
DT   01-NOV-1990 (Rel. 16, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   L-selectin precursor (Lymph node homing receptor) (Leukocyte adhesion
DE   molecule-1) (LAM-1) (LY-22) (Lymphocyte surface MEL-14 antigen)
DE   (Leukocyte-endothelial cell adhesion molecule 1) (LECAM1) (CD62L).
GN   SELL OR LNHR OR LY-22.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=89162048; PubMed=2646713;
RA   Siegelman M.H., van de Rijn M., Weissman I.L.;
RT   "Mouse lymph node homing receptor cDNA clone encodes a glycoprotein
RT   revealing tandem interaction domains.";
RL   Science 243:1165-1172(1989).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=90263086; PubMed=1693096;
RA   Siegelman M.H., Cheng I.C., Weissman I.L., Wakeland E.K.;
RT   "The mouse lymph node homing receptor is identical with the
RT   lymphocyte cell surface marker Ly-22: role of the EGF domain in
RT   endothelial binding.";
RL   Cell 61:611-622(1990).
RN   [3]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=89168433; PubMed=2647302;
RA   Lasky L.A., Singer M.S., Yednock T.A., Dowbenko D., Fennie C.,
RA   Rodriguez H., Nguyen T., Stachel S., Rosen S.D.;
RT   "Cloning of a lymphocyte homing receptor reveals a lectin domain.";
RL   Cell 56:1045-1055(1989).
RN   [4]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6; TISSUE=Hematopoietic;
RX   MEDLINE=22388257; PubMed=12477932;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length
RT   human and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [5]
RP   SEQUENCE OF 1-360 FROM N.A.
RX   MEDLINE=91169529; PubMed=2004776;
RA   Dowbenko D.J., Diep A., Taylor B.A., Lusis A.J., Lasky L.A.;
RT   "Characterization of the murine homing receptor gene reveals
RT   correspondence between protein domains and coding exons.";
RL   Genomics 9:270-277(1991).
CC   -!- FUNCTION: Cell surface adhesion protein. Mediate the adherence of
CC       lymphocytes to endothelial cells of high endothelial venules in
CC       peripheral lymph nodes.
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC   -!- SIMILARITY: Belongs to the selectin/LECAM family.
CC   -!- SIMILARITY: Contains 1 C-type lectin family domain.
CC   -!- SIMILARITY: Contains 1 EGF-like domain.
CC   -!- SIMILARITY: Contains 2 Sushi (SCR) domains.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
```

```
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; X14772; CAA32880.1; -.
DR   EMBL; M36005; AAA39722.1; -.
DR   EMBL; M36058; AAA39723.1; -.
DR   EMBL; M25324; AAA39431.1; -.
DR   EMBL; M64549; AAA75651.1; -.
DR   EMBL; M64440; AAA75651.1; JOINED.
DR   EMBL; M64545; AAA75651.1; JOINED.
DR   EMBL; M64548; AAA75651.1; JOINED.
DR   EMBL; BC052681; AAH52681.1; -.
DR   PIR; A32375; A32375.
DR   HSSP; P14151; 1KJB.
DR   MGD; MGI:98279; Sell.
DR   InterPro; IPR006209; EGF_like.
DR   InterPro; IPR006210; IEGF.
DR   InterPro; IPR001304; Lectin_C.
DR   InterPro; IPR002396; Selectin.
DR   InterPro; IPR000436; Sushi_SCR_CCP.
DR   Pfam; PF00008; EGF; 1.
DR   Pfam; PF00059; lectin_c; 1.
DR   Pfam; PF00084; sushi; 2.
DR   PRINTS; PR00343; SELECTIN.
DR   SMART; SM00032; CCP; 2.
DR   SMART; SM00034; CLECT; 1.
DR   SMART; SM00181; EGF; 1.
DR   PROSITE; PS00615; C_TYPE_LECTIN_1; 1.
DR   PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
DR   PROSITE; PS00022; EGF_1; 1.
DR   PROSITE; PS01186; EGF_2; 1.
DR   PROSITE; PS50026; EGF_3; 1.
KW   Cell adhesion; Transmembrane; Glycoprotein; EGF-like domain; Lectin;
KW   Selectin; Signal; Sushi; Repeat.
FT   SIGNAL        1     28
FT   PROPEP       29     38
FT   CHAIN        39    372       L-SELECTIN.
FT   DOMAIN       39    332       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    333    355       POTENTIAL.
FT   DOMAIN      356    372       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN       55    155       C-TYPE LECTIN (SHORT FORM).
FT   DOMAIN      156    192       EGF-LIKE.
FT   DOMAIN      196    255       SUSHI 1.
FT   DOMAIN      258    317       SUSHI 2.
FT   DISULFID     57    155       BY SIMILARITY.
FT   DISULFID    128    147       BY SIMILARITY.
FT   DISULFID    160    171       BY SIMILARITY.
FT   DISULFID    165    180       BY SIMILARITY.
FT   DISULFID    182    191       BY SIMILARITY.
FT   DISULFID    197    241       BY SIMILARITY.
FT   DISULFID    227    254       BY SIMILARITY.
FT   DISULFID    259    303       BY SIMILARITY.
FT   DISULFID    289    316       BY SIMILARITY.
FT   CARBOHYD     60     60       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    104    104       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    177    177       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    216    216       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    226    226       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    246    246       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    278    278       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    288    288       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    308    308       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    320    320       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CONFLICT     32     32       I -> T (IN REF. 5).
SQ   SEQUENCE   372 AA;  42288 MW;  4433EDF6E4CB2B78 CRC64;

  Query Match             12.9%;  Score 143.5;  DB 1;  Length 372;
  Best Local Similarity   26.5%;  Pred. No. 1.7e-05;
  Matches   41;  Conservative   37;  Mismatches   50;  Indels   27;  Gaps    8;

Qy         26 FHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLP-SDGDFWIGLRRREE 84
              :| : : :|:| |:: |::: ||:|::: |  || ::|| || |  :|||:|:
Db         41 YHYSEKPMNWENARKFCKQNYTDLVAIQNKRE---IE-YLENTLPKSPYYYWIGIRK--- 93

Qy         85 KQSNSTACQDLYAW--TDGSIS-QFRNWYVDEPS--CGSEVCVVMYHQPSAPAGIGGPYM 139
                        :: |  |: ::: :  ||   ||:     | || :| :    :|
Db         94 -------IGKMWTWVGTNKTLTKEAENWGAGEPNNKKSKEDCVEIYIKRERDSG------ 140

Qy        140 FQWNDDRCNMKNNFICKYSDEKPAVPSREAEGEET 174
                :|||| |: :   :|  :  :|   :   |  ||
Db        141 -KWNDDACHKRKAALCYTASCQPGSCNGRGECVET 174


RESULT 31
FCE2_MOUSE
ID   FCE2_MOUSE     STANDARD;      PRT;   331 AA.
AC   P20693; Q61556; Q61557;
DT   01-FEB-1991 (Rel. 17, Created)
DT   01-FEB-1991 (Rel. 17, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   Low affinity immunoglobulin epsilon FC receptor (Lymphocyte IGE
DE   receptor) (FC-epsilon-RII) (CD23).
GN   FCER2 OR FCER2A.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=90017519; PubMed=2529542;
RA   Bettler B., Hofstetter H., Rao M., Yokoyama W.M., Kilchherr F.,
RA   Conrad D.H.;
RT   "Molecular structure and expression of the murine lymphocyte low-
RT   affinity receptor for IgE (Fc epsilon RII).";
RL   Proc. Natl. Acad. Sci. U.S.A. 86:7566-7570(1989).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=90171598; PubMed=2137845;
RA   Gollnick S.O., Trounstine M.L., Yamashita L.C., Kehry M.R.,
RA   Moore K.W.;
RT   "Isolation, characterization, and expression of cDNA clones encoding
RT   the mouse Fc receptor for IgE (Fc epsilon RII)1.";
RL   J. Immunol. 144:1974-1982(1990).
RN   [3]
RP   SEQUENCE FROM N.A. (ISOFORMS B AND C).
RC   STRAIN=DBA/2;
RX   MEDLINE=94372613; PubMed=8086828;
RA   Kondo H., Ichikawa Y., Nakamura K., Tsuchiya S.;
RT   "Cloning of cDNAs for new subtypes of murine low-affinity Fc receptor
RT   for IgE (Fc epsilon RII/CD23).";
RL   Int. Arch. Allergy Immunol. 105:38-48(1994).
RN   [4]
RP   3D-STRUCTURE MODELING OF LECTIN DOMAIN.
RX   MEDLINE=94191542; PubMed=8142907;
RA   Padlan E.A., Helm B.A.;
RT   "Modeling of the lectin-homology domains of the human and murine low-
RT   affinity Fc epsilon receptor (Fc epsilon RII/CD23).";
RL   Receptor 3:325-341(1993).
CC   -!- FUNCTION: This receptor has essential roles in the regulation of
CC       IgE production and in the differentiation of B-cells (it is a B-
CC       cell-specific antigen).
CC   -!- SUBCELLULAR LOCATION: Type II membrane protein.
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=3;
CC       Name=A;
CC         IsoId=P20693-1; Sequence=Displayed;
CC       Name=B;
CC         IsoId=P20693-2; Sequence=VSP_003058;
CC       Name=C;
CC         IsoId=P20693-3; Sequence=VSP_003059;
CC   -!- PTM: N- and O-glycosylated (By similarity).
CC   -!- MISCELLANEOUS: There are two kinds of Fc receptors for IgE, which
CC       differ in both structure and function: high affinity receptors on
CC       basophils and mast cells and low affinity receptors on lymphocytes
CC       and monocytes.
CC   -!- SIMILARITY: Contains 1 C-type lectin family domain.
```

```
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M99371; AAA74898.1; -.
DR   EMBL; M34163; AAA37603.1; -.
DR   EMBL; X64223; CAA45532.1; -.
DR   EMBL; X64224; CAA45533.1; -.
DR   PIR; A43518; LNMSER.
DR   PDB; 1HLJ; 31-JAN-94.
DR   MGD; MGI:95497; Fcer2a.
DR   InterPro; IPR002353; AntifreezeII.
DR   InterPro; IPR001304; Lectin_C.
DR   Pfam; PF00059; lectin_c; 1.
DR   PRINTS; PR00356; ANTIFREEZEII.
DR   SMART; SM00034; CLECT; 1.
DR   PROSITE; PS00615; C_TYPE_LECTIN_1; 1.
DR   PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
KW   Receptor; Antigen; IgE-binding protein; Repeat; Signal-anchor;
KW   Transmembrane; Lectin; Glycoprotein; Alternative splicing;
KW   3D-structure.
FT   DOMAIN        1     23       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     24     49       SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
FT                                 (POTENTIAL).
FT   DOMAIN       50    331       EXTRACELLULAR (POTENTIAL).
FT   DOMAIN      185    298       C-TYPE LECTIN (LONG FORM).
FT   REPEAT       71     91
FT   REPEAT       92    112
FT   REPEAT      113    133
FT   DISULFID    183    311       BY SIMILARITY.
FT   DISULFID    186    197       BY SIMILARITY.
FT   DISULFID    214    305       BY SIMILARITY.
FT   DISULFID    282    296       BY SIMILARITY.
FT   CARBOHYD     65     65       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    114    114       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   VARSPLIC      1      7       MEENEYS -> MNSQNQ (in isoform B).
FT                                 /FTId=VSP_003058.
FT   VARSPLIC      1      7       MEENEYS -> MDTHHT (in isoform C).
FT                                 /FTId=VSP_003059.
FT   STRAND      197    200
FT   HELIX       207    216
FT   TURN        217    218
FT   STRAND      220    221
FT   HELIX       227    237
FT   TURN        238    239
FT   STRAND      242    247
FT   TURN        249    250
FT   TURN        252    253
FT   STRAND      256    257
FT   TURN        258    259
FT   STRAND      260    262
FT   STRAND      268    268
FT   TURN        270    271
FT   TURN        277    278
FT   STRAND      282    285
FT   TURN        287    288
FT   STRAND      291    294
FT   TURN        296    297
FT   STRAND      304    307
SQ   SEQUENCE   331 AA;  37647 MW;  B8C6D65F34ACCDB2 CRC64;

  Query Match             12.7%;  Score 141.5;  DB 1;  Length 331;
  Best Local Similarity   31.8%;  Pred. No. 2.2e-05;
  Matches   47;  Conservative   21;  Mismatches   37;  Indels   43;  Gaps    9;

Qy         17 QRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFW 76
              |: ||   ||   |::  : :|: ||    |:|||| |: ||  : : |     :  | |
Db        194 QQKCY---YFGKGSKQ--WIQARFACSDLQGRLVSIHSQKEQDFLMQHI-----NKKDSW 243

Qy         77 IGLRRREEKQSNSTACQDL-----YAWTDGSISQFRNWYVDEPSCG--SEVCVVMYHQPS 129
              |||              |||     : |:|||   : ||   ||: |   | ||:|
Db        244 IGL-------------QDLNMEGEFVWSDGSPVGYSNWNPGEPNNGGQGEDCVMM----- 285

Qy        130 APAGIGGPYMFQWNDDRC-NMKNNFICK 156
                | |     ||||  | :  : ::|:
Db        286 --RGSG-----QWNDAFCRSYLDAWVCE 306


RESULT 32
LEM2_RAT
ID   LEM2_RAT       STANDARD;      PRT;   549 AA.
AC   P98105;
DT   01-FEB-1996 (Rel. 33, Created)
DT   01-FEB-1996 (Rel. 33, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   E-selectin precursor (Endothelial leukocyte adhesion molecule 1)
DE   (ELAM-1) (Leukocyte-endothelial cell adhesion molecule 2) (LECAM2).
GN   SELE OR ELAM-1.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Lung;
RA   Rosenbloom C.L., Auchampach J.A., Anderson D.C., Manning A.M.;
RL   Submitted (NOV-1993) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: Expressed on cytokine induced endothelial cells and
CC       mediates their binding to Leukocytes. The ligand recognized by
CC       ELAM-1 is sialyl-lewis X (alpha(1->3)fucosylated derivatives of
CC       polylactosamine that are found at the nonreducing termini of
CC       glycolipids).
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC   -!- SIMILARITY: Belongs to the selectin/LECAM family.
CC   -!- SIMILARITY: Contains 1 C-type lectin family domain.
CC   -!- SIMILARITY: Contains 1 EGF-like domain.
CC   -!- SIMILARITY: Contains 5 Sushi (SCR) domains.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; L25527; AAA41113.1; -.
DR   HSSP; P16581; 1KJA.
DR   InterPro; IPR006209; EGF_like.
DR   InterPro; IPR006210; IEGF.
DR   InterPro; IPR001304; Lectin_C.
DR   InterPro; IPR002396; Selectin.
DR   InterPro; IPR000436; Sushi_SCR_CCP.
DR   Pfam; PF00008; EGF; 1.
DR   Pfam; PF00059; lectin_c; 1.
DR   Pfam; PF00084; sushi; 5.
DR   PRINTS; PR00343; SELECTIN.
DR   SMART; SM00032; CCP; 5.
DR   SMART; SM00034; CLECT; 1.
DR   SMART; SM00181; EGF; 1.
DR   PROSITE; PS00615; C_TYPE_LECTIN_1; 1.
DR   PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
DR   PROSITE; PS00022; EGF_1; 1.
DR   PROSITE; PS01186; EGF_2; 1.
DR   PROSITE; PS50026; EGF_3; 1.
KW   Cell adhesion; Transmembrane; Glycoprotein; EGF-like domain; Lectin;
KW   Selectin; Signal; Sushi; Repeat.
FT   SIGNAL        1     21       POTENTIAL.
FT   CHAIN        22    549       E-SELECTIN.
FT   DOMAIN       22    494       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    495    516       POTENTIAL.
```

```
FT   DOMAIN      517    549       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN       22    139       C-TYPE LECTIN.
FT   DOMAIN      140    175       EGF-LIKE.
FT   DOMAIN      179    239       SUSHI 1.
FT   DOMAIN      242    301       SUSHI 2.
FT   DOMAIN      304    364       SUSHI 3.
FT   DOMAIN      367    427       SUSHI 4.
FT   DOMAIN      430    486       SUSHI 5.
FT   DISULFID     40    138       BY SIMILARITY.
FT   DISULFID    111    130       BY SIMILARITY.
FT   DISULFID    143    154       BY SIMILARITY.
FT   DISULFID    148    163       BY SIMILARITY.
FT   DISULFID    165    174       BY SIMILARITY.
FT   DISULFID    180    225       BY SIMILARITY.
FT   DISULFID    210    238       BY SIMILARITY.
FT   DISULFID    243    287       BY SIMILARITY.
FT   DISULFID    273    300       BY SIMILARITY.
FT   DISULFID    305    350       BY SIMILARITY.
FT   DISULFID    336    363       BY SIMILARITY.
FT   DISULFID    368    413       BY SIMILARITY.
FT   DISULFID    399    426       BY SIMILARITY.
FT   DISULFID    431    472       BY SIMILARITY.
FT   DISULFID    458    485       BY SIMILARITY.
FT   CARBOHYD     25     25       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     60     60       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    145    145       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    192    192       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    203    203       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    266    266       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    313    313       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    320    320       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    333    333       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    441    441       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    465    465       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   549 AA;  60079 MW;  85CEECDB7B0144C8 CRC64;

  Query Match             12.7%;  Score 141.5;  DB 1;  Length 549;
  Best Local Similarity   26.5%;  Pred. No. 4.2e-05;
  Matches   36;  Conservative   28;  Mismatches   47;  Indels   25;  Gaps    5;

Qy         25 YFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFWIGLRRREE 84
              |:: :|  : ::||   |:||   ||:|::::|   |       |  |   :|||:|:
Db         23 YYNASSELMTYDEASAYCQRDYTHLVAIQNKEE---INYLNSTLRYSPSYYWIGIRK--- 76

Qy         85 KQSNSTACQDLYAWTDGS---ISQFRNWYVDEPS--CGSEVCVVMYHQPSAPAGIGGPYM 139
                     ::: |         : :||   ||:    :| || :| |    :|:
Db         77 -------VNNVWIWVGTQKPLTEEAKNWAPGEPNNKQRNEDCVEIYIQRPKDSGM----- 124

Qy        140 FQWNDDRCNMKNNFIC 155
               |||:||: |   :|
Db        125 --WNDERCDKKKLALC 138


RESULT 33
LEC3_MEGRO
ID   LEC3_MEGRO     STANDARD;      PRT;   162 AA.
AC   P07439;
DT   01-APR-1988 (Rel. 07, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Lectin BRA-3 precursor.
OS   Megabalanus rosa (Acorn barnacle).
OC   Eukaryota; Metazoa; Arthropoda; Crustacea; Maxillopoda; Cirripedia;
OC   Thoracica; Sessilia; Balanomorpha; Balanoidea; Balanidae;
OC   Megabalanus.
OX   NCBI_TaxID=6680;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=93292994; PubMed=8514190;
RA   Takamatsu N., Takeda T., Kojima M., Heishi M., Muramoto K.,
RA   Kamiya H., Shiba T.;
RT   "Acorn barnacle Megabalanus rosa lectin (BRA-3): cDNA cloning, gene
RT   structure and seasonal changes of mRNA and protein levels.";
RL   Gene 128:251-255(1993).
RN   [2]
RP   SEQUENCE OF 25-162.
RA   Muramoto K., Kamiya H.;
RT   "The amino-acid sequence of a lectin of the acorn barnacle
RT   Megabalanus rosa.";
RL   Biochim. Biophys. Acta 874:285-295(1986).
CC   -!- FUNCTION: Sugar-binding protein which recognizes specific
CC       carbohydrate structures and agglutinate a variety of animal cells
CC       by binding to cell-surface glycoproteins and glycolipids. Calcium-
CC       dependent lectin. Invertebrate lectins may be involved in defense
CC       functions.
CC   -!- SUBUNIT: Homotetramer; disulfide-linked.
CC   -!- TISSUE SPECIFICITY: Coelemic fluid.
CC   -!- MISCELLANEOUS: This lectin binds galactose.
CC   -!- SIMILARITY: Contains 1 C-type lectin family domain.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; D13302; BAA02556.1; -.
DR   EMBL; D13299; BAA02556.1; JOINED.
DR   EMBL; D13300; BAA02556.1; JOINED.
DR   EMBL; D13301; BAA02556.1; JOINED.
DR   HSSP; P23806; 1IXX.
DR   InterPro; IPR002353; AntifreezeII.
DR   InterPro; IPR001304; Lectin_C.
DR   InterPro; IPR003990; Pancreatis_ac.
DR   Pfam; PF00059; lectin_c; 1.
DR   PRINTS; PR00356; ANTIFREEZEII.
DR   PRINTS; PR01504; PNCREATITSAP.
DR   SMART; SM00034; CLECT; 1.
DR   PROSITE; PS00615; C_TYPE_LECTIN_1; 1.
DR   PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
KW   Lectin; Calcium; Signal.
FT   SIGNAL        1     24
FT   CHAIN        25    162       LECTIN BRA-3.
FT   DOMAIN       25    152       C-TYPE LECTIN (LONG FORM).
FT   DISULFID     26     39
FT   DISULFID     56    150
FT   DISULFID    125    142
FT   DISULFID    157    157       INTERCHAIN (WITH C-136 IN OTHER CHAIN).
FT   DISULFID    160    160       INTERCHAIN (WITH C-133 IN OTHER CHAIN).
FT   VARIANT     146    146       K -> R.
SQ   SEQUENCE   162 AA;  18328 MW;  EB7F14E91DD1CB81 CRC64;

  Query Match             12.6%;  Score 140.5;  DB 1;  Length 162;
  Best Local Similarity   24.4%;  Pred. No. 1.1e-05;
  Matches   39;  Conservative   26;  Mismatches   60;  Indels   35;  Gaps    6;

Qy          5 LLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACR--RDGGQLVSIESEDEQKLIE 62
              : : :  | |       |    :  :: :: : :|: ||:    |  | :|:|: |   |
Db         19 ITTAECTCPGNLDWQEYDGHCYWASTYQVRWNDAQLACQTVHPGAYLATIQSQLENAFIS 78

Qy         63 KFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEPS------C 116
              : : |        ||||       |    :  | |::|  : |  |  : |:      |
Db         79 ETVSN-----NRLWIGL--------NDIDLEGHYVWSNGEATDFTYWSSNNPNNWENQDC 125

Qy        117 GSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICK 156
              |    || |   :           ||:|| ||   ||:||
Db        126 G----VVNYDTVTG----------QWDDDDCNKNKNFLCK 151


RESULT 34
LEM1_BOVIN
ID   LEM1_BOVIN     STANDARD;      PRT;   370 AA.
AC   P98131;
DT   01-FEB-1996 (Rel. 33, Created)
```

```
DT   01-FEB-1996 (Rel. 33, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   L-selectin precursor (Lymph node homing receptor) (Leukocyte adhesion
DE   molecule-1) (LAM-1) (Leukocyte-endothelial cell adhesion molecule 1)
DE   (LECAM1) (CD62L).
GN   SELL.
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=92164727; PubMed=1371468;
RA   Walcheck B., White M., Kurk S., Kishimoto T.K., Jutila M.A.;
RT   "Characterization of the bovine peripheral lymph node homing
RT   receptor: a lectin cell adhesion molecule (LECAM).";
RL   Eur. J. Immunol. 22:469-476(1992).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=94055053; PubMed=7694420;
RA   Bosworth B.T., Dowbenko D., Shuster D.E., Harp J.A.;
RT   "Bovine L-selectin: a peripheral lymphocyte homing receptor.";
RL   Vet. Immunol. Immunopathol. 37:201-215(1993).
CC   -!- FUNCTION: Cell surface adhesion protein. Mediate the adherence of
CC       lymphocytes to endothelial cells of high endothelial venules in
CC       peripheral lymph nodes.
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC   -!- SIMILARITY: Belongs to the selectin/LECAM family.
CC   -!- SIMILARITY: Contains 1 C-type lectin family domain.
CC   -!- SIMILARITY: Contains 1 EGF-like domain.
CC   -!- SIMILARITY: Contains 2 Sushi (SCR) domains.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X62882; CAA44676.1; -.
DR   PIR; S22124; S22124.
DR   HSSP; P14151; 1KJB.
DR   InterPro; IPR006209; EGF_like.
DR   InterPro; IPR006210; IEGF.
DR   InterPro; IPR001304; Lectin_C.
DR   InterPro; IPR002396; Selectin.
DR   InterPro; IPR000436; Sushi_SCR_CCP.
DR   Pfam; PF00008; EGF; 1.
DR   Pfam; PF00059; lectin_c; 1.
DR   Pfam; PF00084; sushi; 2.
DR   PRINTS; PR00343; SELECTIN.
DR   SMART; SM00032; CCP; 2.
DR   SMART; SM00034; CLECT; 1.
DR   SMART; SM00181; EGF; 1.
DR   PROSITE; PS00615; C_TYPE_LECTIN_1; 1.
DR   PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
DR   PROSITE; PS00022; EGF_1; 1.
DR   PROSITE; PS01186; EGF_2; 1.
DR   PROSITE; PS50026; EGF_3; 1.
KW   Cell adhesion; Transmembrane; Glycoprotein; EGF-like domain; Lectin;
KW   Selectin; Signal; Sushi; Repeat.
FT   SIGNAL       1    28       POTENTIAL.
FT   PROPEP      29    38       POTENTIAL.
FT   CHAIN       39   370       L-SELECTIN.
FT   DOMAIN      39   333       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   334   354       POTENTIAL.
FT   DOMAIN     355   370       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN      55   155       C-TYPE LECTIN (SHORT FORM).
FT   DOMAIN     156   192       EGF-LIKE.
FT   DOMAIN     196   255       SUSHI 1.
FT   DOMAIN     258   317       SUSHI 2.
FT   DISULFID    57   155       BY SIMILARITY.
FT   DISULFID   128   147       BY SIMILARITY.
FT   DISULFID   160   171       BY SIMILARITY.
FT   DISULFID   165   180       BY SIMILARITY.
FT   DISULFID   182   191       BY SIMILARITY.
FT   DISULFID   197   241       BY SIMILARITY.
FT   DISULFID   227   254       BY SIMILARITY.
FT   DISULFID   259   303       BY SIMILARITY.
FT   DISULFID   289   316       BY SIMILARITY.
FT   CARBOHYD    60    60       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    77    77       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   104   104       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   177   177       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   216   216       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   226   226       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   246   246       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   308   308       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   320   320       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   370 AA;  41971 MW;  92168F8116AE9228 CRC64;

  Query Match          12.6%;  Score 140;  DB 1;  Length 370;
  Best Local Similarity  25.3%;  Pred. No. 3.5e-05;
  Matches   46;  Conservative   35;  Mismatches   57;  Indels   44;  Gaps    9;

Qy        26 FHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLP-SDGDFWIGLRRREE 84
             :| : | : :|:|:  || :   ||:|::: |   || ::   || |   :|||:|: |
Db        41 YHYSKRPMPWEKARAFCRENYTDLVAIQNKGE---IE-YLNKTLPFSRTYYWIGIRKVE- 95

Qy        85 KQSNSTACQDLYAW--TDGSIS-QFRNWYVDEPS--CGSEVCVVMYHQPSAPAGIGGPYM 139
                    :: |  |: |:: : :||   ||:      | || :| : :  :|
Db        96 ---------GVWTWVGTNKSLTEEAKNWGAGEPNNRKSKEDCVEIYIKRNKDSG------ 140

Qy       140 FQWNDDRCNMKNNFICKYSDEKPAVPSREAE----------------GEETELTTPVLP 182
              :|||| |:     :|  :  ||   |   :                | | :  |  :|
Db       141 -KWNDDACHKAKTALCYTASCKPWSCSGHGQCVEVINNYTCNCDLGYYGPECQFVTQCVP 199

Qy       183 EE 184
              |
Db       200 LE 201


RESULT 35
PSPA_HUMAN
ID   PSPA_HUMAN     STANDARD;      PRT;   248 AA.
AC   P07714;
DT   01-APR-1988 (Rel. 07, Created)
DT   01-APR-1990 (Rel. 14, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Pulmonary surfactant-associated protein A precursor (SP-A) (PSP-A)
DE   (PSAP) (Alveolar proteinosis protein) (35 kDa pulmonary surfactant-
DE   associated protein).
GN   (SFTPA1 OR SFTPA OR SFTP1 OR PSAP) AND (SFTPA2 OR SFTPA).
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=86250832; PubMed=3755136;
RA   Floros J., Steinbrink R., Jacobs K., Phelps D., Kriz R., Recny M.,
RA   Sultzman L., Jones S., Taeusch H.W., Frank H.A., Fritsch E.F.;
RT   "Isolation and characterization of cDNA clones for the 35-kDa
RT   pulmonary surfactant-associated protein.";
RL   J. Biol. Chem. 261:9029-9033(1986).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=86014366; PubMed=2995821;
RA   White R.T., Damm D., Miller J., Spratt K., Schilling J., Hawgood S.,
RA   Benson B., Cordell B.;
RT   "Isolation and characterization of the human pulmonary surfactant
RT   apoprotein gene.";
RL   Nature 317:361-363(1985).
RN   [3]
```

```
RP   SEQUENCE FROM N.A.
RX   MEDLINE=92198680; PubMed=1372511;
RA   Katyal S.L., Singh G., Locker J.L.;
RT   "Characterization of a second human pulmonary surfactant-associated
RT   protein SP-A gene.";
RL   Am. J. Respir. Cell Mol. Biol. 6:446-452(1992).
CC   -!- FUNCTION: In presence of calcium ions, PSAP binds to surfactant
CC       phospholipids and contributes to lower the surface tension at the
CC       air-liquid interface in the alveoli of the mammalian lung and is
CC       essential for normal respiration.
CC   -!- SUBUNIT: Oligomeric complex of 6 set of homotrimers.
CC   -!- SUBCELLULAR LOCATION: Extracellular.
CC   -!- MISCELLANEOUS: Pulmonary surfactant consists of 90% lipid and 10%
CC       protein. There are 4 surfactant-associated protein: 2 collagenous,
CC       carbohydrate-binding glycoproteins (SP-A and SP-D) and 2 small
CC       hydrophobic proteins (SP-B and SP-C).
CC   -!- SIMILARITY: Contains 1 collagenous domain.
CC   -!- SIMILARITY: Contains 1 C-type lectin family domain.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; M13686; AAA60211.1; -.
DR   EMBL; K03475; AAA36520.1; -.
DR   EMBL; M30838; AAA36510.1; -.
DR   EMBL; M68519; AAA60319.1; -.
DR   PIR; A25720; LNHUP6.
DR   HSSP; P22897; 1EGG.
DR   Genew; HGNC:10798; SFTPA1.
DR   Genew; HGNC:10799; SFTPA2.
DR   MIM; 178630; -.
DR   MIM; 178642; -.
DR   GO; GO:0005319; F:lipid transporter activity; TAS.
DR   InterPro; IPR008160; Collagen.
DR   InterPro; IPR001304; Lectin_C.
DR   Pfam; PF01391; Collagen; 2.
DR   Pfam; PF00059; lectin_c; 1.
DR   SMART; SM00034; CLECT; 1.
DR   PROSITE; PS00615; C_TYPE_LECTIN_1; 1.
DR   PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
KW   Glycoprotein; Calcium; Surface film; Gaseous exchange; Hydroxylation;
KW   Signal; Lectin; Collagen; Repeat; Polymorphism.
FT   SIGNAL       1     20
FT   CHAIN       21    248       PULMONARY SURFACTANT-ASSOCIATED PROTEIN
FT                               A.
FT   DOMAIN      28    100       COLLAGEN-LIKE.
FT   DOMAIN     153    248       C-TYPE LECTIN (SHORT FORM).
FT   DISULFID   155    246       BY SIMILARITY.
FT   DISULFID   224    238       BY SIMILARITY.
FT   CARBOHYD   207    207       N-LINKED (GLCNAC. . .) (PROBABLE).
FT   VARIANT      9      9       N -> T.
FT                               /FTId=VAR_004184.
FT   VARIANT     50     50       L -> V.
FT                               /FTId=VAR_012231.
FT   VARIANT     66     66       M -> T.
FT                               /FTId=VAR_004185.
FT   VARIANT     73     73       D -> N.
FT                               /FTId=VAR_004186.
FT   VARIANT     81     81       I -> V.
FT                               /FTId=VAR_004187.
FT   VARIANT     85     85       C -> R.
FT                               /FTId=VAR_004188.
FT   VARIANT    219    219       R -> W.
FT                               /FTId=VAR_012232.
FT   VARIANT    223    223       Q -> K.
FT                               /FTId=VAR_012233.
FT   CONFLICT    19     19       A -> V (IN REF. 2).
FT   CONFLICT    45     45       D -> H (IN REF. 2).
FT   CONFLICT    54     54       P -> L (IN REF. 2).
FT   CONFLICT    91     91       P -> A (IN REF. 3).
FT   CONFLICT   100    100       P -> R (IN REF. 2; AAA36510).
FT   CONFLICT   247    247       E -> D (IN REF. 3).
SQ   SEQUENCE   248 AA;  26214 MW;  6A9F0C3488BF3633 CRC64;
```

```
  Query Match             12.5%;  Score 139;  DB 1;  Length 248;
  Best Local Similarity   25.2%;  Pred. No. 2.6e-05;
  Matches   36;  Conservative   18;  Mismatches   45;  Indels   44;  Gaps    4;

Qy         26 FHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLL----------PSDGDF 75
              |    : : |:  :||| | ||::    : :| : |  |::             || |||
Db        139 FSSNGQSITFDAIQEACARAGGRIAVPRNPEENEAIASFVKKYNTYAYVGLTEGPSPGDF 198

Qy         76 WIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEPS-CGSEVCVVMYHQPSAPAGI 134
                                  ::||:   : |||  ||:  | | || ||
Db        199 --------------------RYSDGTPVNYTNWYRGEPAGRGKEQCVEMYTDG------ 231

Qy        135 GGPYMFQWNDDRCNMKNNFICKY 157
                 ||||  |      ||::
Db        232 ------QWNDRNCLYSRLTICEF 248
```

```
RESULT 36
LECA_SARPE
ID   LECA_SARPE     STANDARD;      PRT;   283 AA.
AC   P05047;
DT   13-AUG-1987 (Rel. 05, Created)
DT   13-AUG-1987 (Rel. 05, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Lectin, alpha subunit precursor.
OS   Sarcophaga peregrina (Flesh fly) (Boettcherisca peregrina).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Oestroidea;
OC   Sarcophagidae; Sarcophaga.
OX   NCBI_TaxID=7386;
RN   [1]
RP   SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX   MEDLINE=86008294; PubMed=2413021;
RA   Takahashi H., Komano H., Kawaguchi N., Kitamura N., Nakanishi S.,
RA   Natori S.;
RT   "Cloning and sequencing of cDNA of Sarcophaga peregrina humoral
RT   lectin induced on injury of the body wall.";
RL   J. Biol. Chem. 260:12228-12233(1985).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=94139722; PubMed=8307011;
RA   Matsui M., Kobayashi A., Kubo T., Natori S.;
RT   "Purification and characterization of ATBP, a novel protein that
RT   binds to A/T stretches in three segments of the Sarcophaga lectin
RT   gene.";
RL   Eur. J. Biochem. 219:449-454(1994).
RN   [3]
RP   SEQUENCE OF 1-38 FROM N.A.
RX   MEDLINE=90089397; PubMed=2480809;
RA   Kobayashi A., Hirai H., Kubo T., Veno K., Nakanishi Y., Natori S.;
RT   "Cloning and in vitro transcription of the Sarcophaga lectin gene.";
RL   Biochim. Biophys. Acta 1009:244-250(1989).
CC   -!- FUNCTION: Role in the defence system of the organism against
CC       microorganisms. This lectin binds galactose.
CC   -!- INDUCTION: By injury of the body wall.
CC   -!- SIMILARITY: Contains 1 C-type lectin family domain.
CC   -!- CAUTION: It is uncertain whether Met-1 or Met-5 is the initiator.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; M11673; AAA29983.1; -.
```

```
DR   EMBL; D14870; BAA03586.1; -.
DR   EMBL; X16659; CAA34645.1; -.
DR   PIR; S07759; LNFHLS.
DR   InterPro; IPR001304; Lectin_C.
DR   Pfam; PF00059; lectin_c; 1.
DR   SMART; SM00034; CLECT; 1.
DR   PROSITE; PS00615; C_TYPE_LECTIN_1; 1.
DR   PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
KW   Lectin; Signal.
FT   SIGNAL        1     23
FT   CHAIN        24    283       LECTIN, ALPHA SUBUNIT.
FT   DOMAIN       51    159       C-TYPE LECTIN (SHORT FORM).
FT   DISULFID     53    157       BY SIMILARITY.
FT   DISULFID    132    149       BY SIMILARITY.
SQ   SEQUENCE   283 AA;  32991 MW;  F324BF1A1140B3AC CRC64;

  Query Match             12.5%;  Score 139;  DB 1;  Length 283;
  Best Local Similarity   23.4%;  Pred. No. 3.1e-05;
  Matches   41;  Conservative   39;  Mismatches   79;  Indels   16;  Gaps    6;

Qy         25 YFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFWIGLRRREE 84
              |  :|  : |: :|    | |    |||:||| |:   |   :: ::    : |:|     :
Db         36 YLIETELKYNWHQAWHECARHDQQLVTIESADKNNAIIDLVKRVVGKSHNLWLG---GND 92

Qy         85 KQSNSTACQDLYAWT-DGSISQFRNWYVDEPS--CGSEVCVVMYHQPSAPAGIGGPYMFQ 141
              : |:|      : |: |    | |  : |      | || ::   : |       ::|
Db         93 EYSSSRDYGRPFFWSPTGQAFSFAYWSENNPDNYKHQEHCVHIW--DTKP-------LYQ 143

Qy        142 WNDDRCNMKNNFICKYSDEKPAV-PSREAEGEETELTTPVLPEETQEEDAKKTFK 195
              |||: ||:|  :||: :  :     : : : |  ::|   :  |  :  ||:: :
Db        144 WNDNDCNVKMGYICEPNHFRETYDQALKQKCEAIKITNSKISTEFDQLHAKQSLE 198


RESULT 37
IXA_TRIFL
ID   IXA_TRIFL      STANDARD;      PRT;   152 AA.
AC   P23806; Q91246;
DT   01-NOV-1991 (Rel. 20, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Coagulation factor IX/factor X-binding protein A chain precursor
DE   (IX/X-BP).
OS   Trimeresurus flavoviridis (Habu).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
OC   Viperidae; Crotalinae; Trimeresurus.
OX   NCBI_TaxID=88087;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=96184662; PubMed=8645314;
RA   Matsuzaki R., Yoshihara E., Yamada M., Shima K., Atoda H., Morita T.;
RT   "cDNA cloning of IX/X-BP, a heterogeneous two-chain anticoagulant
RT   protein from snake venom.";
RL   Biochem. Biophys. Res. Commun. 220:382-387(1996).
RN   [2]
RP   SEQUENCE OF 24-152.
RC   TISSUE=Venom;
RX   MEDLINE=91332000; PubMed=1831197;
RA   Atoda H., Hyuga M., Morita T.;
RT   "The primary structure of coagulation factor IX/factor X-binding
RT   protein isolated from the venom of Trimeresurus flavoviridis.
RT   Homology with asialoglycoprotein receptors, proteoglycan core
RT   protein, tetranectin, and lymphocyte Fc epsilon receptor for
RT   immunoglobulin E.";
RL   J. Biol. Chem. 266:14903-14911(1991).
RN   [3]
RP   X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS).
RX   MEDLINE=97331317; PubMed=9187649;
RA   Mizuno H., Fujimoto Z., Koizumi M., Kano H., Atoda H., Morita T.;
RT   "Structure of coagulation factors IX/X-binding protein, a heterodimer
RT   of C-type lectin domains.";
RL   Nat. Struct. Biol. 4:438-441(1997).
CC   -!- FUNCTION: Anticoagulant protein which binds with factor IX and
CC       factor X in the presence of calcium with a 1 to 1 stoichiometry.
CC   -!- SUBUNIT: Heterodimer of chains A and B; disulfide-linked.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- MISCELLANEOUS: Calcium is required for ligand binding.
CC   -!- SIMILARITY: Contains 1 C-type lectin family domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; D83331; BAA11887.1; -.
DR   PIR; JC4690; JC4690.
DR   PDB; 1IXX; 06-MAY-98.
DR   InterPro; IPR001304; Lectin_C.
DR   InterPro; IPR003990; Pancreatis_ac.
DR   Pfam; PF00059; lectin_c; 1.
DR   PRINTS; PR01504; PNCREATITSAP.
DR   SMART; SM00034; CLECT; 1.
DR   PROSITE; PS00615; C_TYPE_LECTIN_1; 1.
DR   PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
KW   Lectin; Calcium; Signal; 3D-structure.
FT   SIGNAL        1     23
FT   CHAIN        24    152       COAGULATION FACTOR IX/FACTOR X-BINDING
FT                                 PROTEIN A CHAIN.
FT   DOMAIN       24    152       C-TYPE LECTIN (LONG FORM).
FT   DISULFID     25     36       BY SIMILARITY.
FT   DISULFID     53    150       BY SIMILARITY.
FT   DISULFID    102    102       INTERCHAIN (WITH C-98 IN B CHAIN).
FT   DISULFID    125    142       BY SIMILARITY.
FT   TURN         27     28
FT   STRAND       30     31
FT   STRAND       36     44
FT   HELIX        46     56
FT   TURN         58     59
FT   STRAND       61     62
FT   HELIX        68     81
FT   STRAND       89     95
FT   STRAND      105    105
FT   TURN        107    108
FT   STRAND      111    111
FT   STRAND      117    117
FT   HELIX       119    121
FT   STRAND      125    128
FT   HELIX       130    132
FT   TURN        133    134
FT   STRAND      137    140
FT   TURN        142    143
FT   STRAND      146    152
SQ   SEQUENCE   152 AA;  17213 MW;  FB3DDD2369009263 CRC64;

  Query Match             12.4%;  Score 138.5;  DB 1;  Length 152;
  Best Local Similarity   28.4%;  Pred. No. 1.6e-05;
  Matches   40;  Conservative   23;  Mismatches   49;  Indels   29;  Gaps    7;

Qy         20 CYKVIYFHDTSRRLNFEEAKEAC--RRDGGQLVSIESEDEQKLIEKFI-ENLLPSDGDFW 76
              |||     : |      :|:|:  |  :  |  ||||||  |    : : : :|:  |  |
Db         36 CYKAFEKYKT-----WEDAERVCTEQAKGAHLVSIESSGEADFVAQLVTQNMKRLDFYIW 90

Qy         77 IGLR-RREEKQSNSTACQDLYAWTDGSISQFRNWYVDEPSCGSEVCVVMYHQPSAPAGIG 135
              |||| : : || ||         |:|||    : ||    |     |: |: :  :
Db         91 IGLRVQGKVKQCNS-------EWSDGSSVSYENWIEAE----SKTCLGLEKETD------ 133

Qy        136 GPYMFQWNDDRCNMKNNFICK 156
                :| :   |  :| |:|:
Db        134 ---FRKWVNIYCGQQNPFVCE 151


RESULT 38
TETN_MOUSE
```

```
ID   TETN_MOUSE     STANDARD;      PRT;   202 AA.
AC   P43025;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-FEB-1996 (Rel. 33, Last sequence update)
DT   15-JUL-1999 (Rel. 38, Last annotation update)
DE   Tetranectin precursor (TN) (Plasminogen-kringle 4 binding protein).
GN   TNA.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6 X CBA; TISSUE=Lung;
RX   MEDLINE=95137396; PubMed=7835708;
RA   Soerensen C.B., Berglund L., Petersen T.E.;
RT   "Cloning of a cDNA encoding murine tetranectin.";
RL   Gene 152:243-245(1995).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BALB/c;
RX   MEDLINE=96116955; PubMed=8563165;
RA   Ibaraki K., Kozak C.A., Wewer U.M., Albrechtsen R., Young M.F.;
RT   "Mouse tetranectin: cDNA sequence, tissue-specific expression, and
RT   chromosomal mapping.";
RL   Mamm. Genome 6:693-696(1995).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BALB/c; TISSUE=Liver;
RX   MEDLINE=98072445; PubMed=9409787;
RA   Soerensen C.B., Berglund L., Petersen T.E.;
RT   "Cloning of the murine tetranectin gene and 5'-flanking region.";
RL   Gene 201:199-202(1997).
CC   -!- FUNCTION: Tetranectin binds to plasminogen and to isolated kringle
CC       4. May be involved in the packaging of molecules destined for
CC       exocytosis (By similarity).
CC   -!- SUBUNIT: Homotrimer (By similarity).
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- TISSUE SPECIFICITY: Highest expression in lung and skeletal
CC       muscle.
CC   -!- SIMILARITY: Contains 1 C-type lectin family domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X79199; CAA55791.1; -.
DR   EMBL; U08595; AAA96811.1; -.
DR   EMBL; X98122; CAA66804.1; -.
DR   PIR; JC4031; JC4031.
DR   HSSP; P05452; 1TN3.
DR   MGD; MGI:104540; Tna.
DR   GO; GO:0001501; P:skeletal development; IMP.
DR   InterPro; IPR001304; Lectin_C.
DR   InterPro; IPR003990; Pancreatis_ac.
DR   Pfam; PF00059; lectin_c; 1.
DR   PRINTS; PR01504; PNCREATITSAP.
DR   SMART; SM00034; CLECT; 1.
DR   PROSITE; PS00615; C_TYPE_LECTIN_1; 1.
DR   PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
KW   Lectin; Plasma; Signal.
FT   SIGNAL        1     21       BY SIMILARITY.
FT   CHAIN        22    202       TETRANECTIN.
FT   DOMAIN       77    198       C-TYPE LECTIN.
FT   DISULFID     71     81       BY SIMILARITY.
FT   DISULFID     98    197       BY SIMILARITY.
FT   DISULFID    173    189       BY SIMILARITY.
FT   CONFLICT     19     20       LT -> VI (IN REF. 2).
FT   CONFLICT     84     84       A -> T (IN REF. 2).
FT   CONFLICT    180    180       A -> R (IN REF. 2).
FT   CONFLICT    188    188       R -> Q (IN REF. 2).
SQ   SEQUENCE   202 AA;  22257 MW;  639E7334D58EB04E CRC64;

  Query Match             12.4%;  Score 138.5;  DB 1;  Length 202;
  Best Local Similarity   26.3%;  Pred. No. 2.3e-05;
  Matches   41;  Conservative   30;  Mismatches   56;  Indels   29;  Gaps    7;

Qy          9 QPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENL 68
              | ||  ||:     :: |   ::   | || | |   || | : :|| | : : ::  :
Db         68 QTVCLKGTKVNLKCLLAF---TQPKTFHEASEDCISQGGTLGTPQSELENEALFEYARHS 124

Qy         69 LPSDGDFWIGLRRREEKQSNSTACQDLYAWTD--GSISQFRNWYVD---EPSCG-SEVCV 122
              : :| : |:||        |  | :   || |  | :  ::||  :    :|  | :| |
Db        125 VGNDANIWLGL--------NDMAAEG--AWVDMTGGLLAYKNWETEITTQPDGGKAENC- 173

Qy        123 VMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYS 158
                      | : |    :| | ||  :  :||:::
Db        174 ---------AALSGAANGKWFDKRCRDQLPYICQFA 200


RESULT 39
LEM1_MACMU
ID   LEM1_MACMU     STANDARD;      PRT;   372 AA.
AC   Q95198;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   L-selectin precursor (Lymph node homing receptor) (Leukocyte adhesion
DE   molecule-1) (LAM-1) (Leukocyte-endothelial cell adhesion molecule 1)
DE   (LECAM1) (CD62L).
GN   SELL.
OS   Macaca mulatta (Rhesus macaque).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
OC   Cercopithecinae; Macaca.
OX   NCBI_TaxID=9544;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Budman J.I., Fu H., Johnson C.E., Thakur A.B., Berg E.L.,
RA   Tsurushita N.;
RL   Submitted (NOV-1996) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: Cell surface adhesion protein. Mediate the adherence of
CC       lymphocytes to endothelial cells of high endothelial venules in
CC       peripheral lymph nodes.
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC   -!- SIMILARITY: Belongs to the selectin/LECAM family.
CC   -!- SIMILARITY: Contains 1 C-type lectin family domain.
CC   -!- SIMILARITY: Contains 1 EGF-like domain.
CC   -!- SIMILARITY: Contains 2 Sushi (SCR) domains.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U73730; AAB18246.1; -.
DR   HSSP; P14151; 1KJB.
DR   InterPro; IPR006209; EGF_like.
DR   InterPro; IPR006210; IEGF.
DR   InterPro; IPR001304; Lectin_C.
DR   InterPro; IPR002396; Selectin.
DR   InterPro; IPR000436; Sushi_SCR_CCP.
DR   Pfam; PF00008; EGF; 1.
DR   Pfam; PF00059; lectin_c; 1.
DR   Pfam; PF00084; sushi; 2.
DR   PRINTS; PR00343; SELECTIN.
DR   SMART; SM00032; CCP; 2.
DR   SMART; SM00034; CLECT; 1.
DR   SMART; SM00181; EGF; 1.
```

```
DR   PROSITE; PS00615; C_TYPE_LECTIN_1; 1.
DR   PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
DR   PROSITE; PS00022; EGF_1; 1.
DR   PROSITE; PS01186; EGF_2; 1.
DR   PROSITE; PS50026; EGF_3; 1.
KW   Cell adhesion; Transmembrane; Glycoprotein; EGF-like domain; Lectin;
KW   Selectin; Signal; Sushi; Repeat.
FT   SIGNAL       1     28       BY SIMILARITY.
FT   PROPEP      29     38       BY SIMILARITY.
FT   CHAIN       39    372       L-SELECTIN.
FT   DOMAIN      39    332       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   333    355       POTENTIAL.
FT   DOMAIN     356    372       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN      55    155       C-TYPE LECTIN (SHORT FORM).
FT   DOMAIN     156    192       EGF-LIKE.
FT   DOMAIN     196    255       SUSHI 1.
FT   DOMAIN     258    317       SUSHI 2.
FT   DISULFID    57    155       BY SIMILARITY.
FT   DISULFID   128    147       BY SIMILARITY.
FT   DISULFID   160    171       BY SIMILARITY.
FT   DISULFID   165    180       BY SIMILARITY.
FT   DISULFID   182    191       BY SIMILARITY.
FT   DISULFID   197    241       BY SIMILARITY.
FT   DISULFID   227    254       BY SIMILARITY.
FT   DISULFID   259    303       BY SIMILARITY.
FT   DISULFID   289    316       BY SIMILARITY.
FT   CARBOHYD    60     60       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   104    104       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   177    177       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   226    226       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   232    232       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   246    246       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   271    271       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   372 AA;  42109 MW;  088D7DD5AC549D6D CRC64;

  Query Match             12.3%;  Score 137.5;  DB 1;  Length 372;
  Best Local Similarity   28.0%;  Pred. No. 5.9e-05;
  Matches   44;  Conservative   31;  Mismatches   53;  Indels   29;  Gaps    9;

Qy         26 FHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLP-SDGDFWIGLRRREE 84
              :| :   :|:::|:  || :   ||:|::: |   || ::|  || |   :|||:|:
Db         41 YHYSENPMNWQKARRFCRENYTDLVAIQNKAE---IE-YLEKTLPFSPSYYWIGIRK--- 93

Qy         85 KQSNSTACQDLYAW--TDGSISQ-FRNWYVDEPS--CGSEVCVVMYHQPSAPAGIGGPYM 139
                      :: |  |: |::|    ||    ||:      | || :| :     ||
Db         94 -------IGGIWTWVGTNKSLTQEAENWGDGEPNNKKNKEDCVEIYIKRKKDAG------ 140

Qy        140 FQWNDDRCNMKNNFICKYSDEKPAVPSREAEGEETEL 176
              :|||| |:     :|  :  :|   |    ||  |:
Db        141 -KWNDDACHKPKAALCYTASCQPW--SCSGHGECVEI 174


RESULT 40
LEM1_PAPHA
ID   LEM1_PAPHA     STANDARD;      PRT;   372 AA.
AC   Q28768;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   L-selectin precursor (Lymph node homing receptor) (Leukocyte adhesion
DE   molecule-1) (LAM-1) (Leukocyte-endothelial cell adhesion molecule 1)
DE   (LECAM1) (CD62L).
GN   SELL.
OS   Papio hamadryas (Hamadryas baboon).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
OC   Cercopithecinae; Papio.
OX   NCBI_TaxID=9557;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=97128794; PubMed=8973334;
RA   Tsurushita N., Fu H., Berg E.L.;
RT   "PCR cloning of the cDNA encoding baboon L-selectin.";
RL   Gene 181:219-220(1996).
CC   -!- FUNCTION: Cell surface adhesion protein. Mediate the adherence of
CC       lymphocytes to endothelial cells of high endothelial venules in
CC       peripheral lymph nodes.
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC   -!- SIMILARITY: Belongs to the selectin/LECAM family.
CC   -!- SIMILARITY: Contains 1 C-type lectin family domain.
CC   -!- SIMILARITY: Contains 1 EGF-like domain.
CC   -!- SIMILARITY: Contains 2 Sushi (SCR) domains.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; U52074; AAB40903.1; -.
DR   PIR; JC5377; JC5377.
DR   HSSP; P14151; 1KJB.
DR   InterPro; IPR006209; EGF_like.
DR   InterPro; IPR006210; IEGF.
DR   InterPro; IPR001304; Lectin_C.
DR   InterPro; IPR002396; Selectin.
DR   InterPro; IPR000436; Sushi_SCR_CCP.
DR   Pfam; PF00008; EGF; 1.
DR   Pfam; PF00059; lectin_c; 1.
DR   Pfam; PF00084; sushi; 2.
DR   PRINTS; PR00343; SELECTIN.
DR   SMART; SM00032; CCP; 2.
DR   SMART; SM00034; CLECT; 1.
DR   SMART; SM00181; EGF; 1.
DR   PROSITE; PS00615; C_TYPE_LECTIN_1; 1.
DR   PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
DR   PROSITE; PS00022; EGF_1; 1.
DR   PROSITE; PS01186; EGF_2; 1.
DR   PROSITE; PS50026; EGF_3; 1.
KW   Cell adhesion; Transmembrane; Glycoprotein; EGF-like domain; Lectin;
KW   Selectin; Signal; Sushi; Repeat.
FT   SIGNAL       1     28       BY SIMILARITY.
FT   PROPEP      29     38       BY SIMILARITY.
FT   CHAIN       39    372       L-SELECTIN.
FT   DOMAIN      39    332       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   333    355       POTENTIAL.
FT   DOMAIN     356    372       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN      55    155       C-TYPE LECTIN (SHORT FORM).
FT   DOMAIN     156    192       EGF-LIKE.
FT   DOMAIN     196    255       SUSHI 1.
FT   DOMAIN     258    317       SUSHI 2.
FT   DISULFID    57    155       BY SIMILARITY.
FT   DISULFID   128    147       BY SIMILARITY.
FT   DISULFID   160    171       BY SIMILARITY.
FT   DISULFID   165    180       BY SIMILARITY.
FT   DISULFID   182    191       BY SIMILARITY.
FT   DISULFID   197    241       BY SIMILARITY.
FT   DISULFID   227    254       BY SIMILARITY.
FT   DISULFID   259    303       BY SIMILARITY.
FT   DISULFID   289    316       BY SIMILARITY.
FT   CARBOHYD    60     60       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   104    104       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   177    177       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   226    226       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   232    232       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   246    246       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   271    271       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   372 AA;  42091 MW;  64E7BDD5AC549D69 CRC64;

  Query Match             12.3%;  Score 137.5;  DB 1;  Length 372;
  Best Local Similarity   28.0%;  Pred. No. 5.9e-05;
  Matches   44;  Conservative   31;  Mismatches   53;  Indels   29;  Gaps    9;

Qy         26 FHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLP-SDGDFWIGLRRREE 84
```

```
             :|  :    :|:::|:  ||  :    ||:|:::  |    ||  ::|   ||  |     :|||:|:
Db        41 YHYSENPMNWQKARRFCRENYTDLVAIQNKAE---IE-YLEKTLPFSPSYYWIGIRK--- 93

Qy        85 KQSNSTACQDLYAW--TDGSISQ-FRNWYVDEPS--CGSEVCVVMYHQPSAPAGIGGPYM 139
                           :: |   |: |::|     ||     ||:        | || :|  :     ||
Db        94 -------IGGIWTWVGTNKSLTQEAENWGDGEPNNKKNKEDCVEIYIKRKKDAG------ 140

Qy       140 FQWNDDRCNMKNNFICKYSDEKPAVPSREAEGEETEL 176
               :||||  |:       :|   :   :|     |      ||   |:
Db       141 -KWNDDACHKPKAALCYTASCQPW--SCSGHGECVEI 174

RESULT 41
LITH_BOVIN
ID   LITH_BOVIN     STANDARD;      PRT;   175 AA.
AC   P23132;
DT   01-NOV-1991 (Rel. 20, Created)
DT   01-NOV-1991 (Rel. 20, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Lithostathine precursor (Pancreatic stone protein) (PSP) (Pancreatic
DE   thread protein) (PTP) (Islet of langerhans regenerating protein) (REG)
DE   (Islet cells regeneration factor) (ICRF).
GN   PTP.
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=90368981; PubMed=2394826;
RA   de la Monte S.M., Ozturk M., Wands J.R.;
RT   "Enhanced expression of an exocrine pancreatic protein in Alzheimer's
RT   disease and the developing human brain.";
RL   J. Clin. Invest. 86:1004-1013(1990).
RN   [2]
RP   SEQUENCE OF 38-138 AND 141-175.
RX   MEDLINE=91197388; PubMed=2085387;
RA   Cai L., Harris W.R., Marshak D.R., Gross J., Crabb J.W.;
RT   "Structural analysis of bovine pancreatic thread protein.";
RL   J. Protein Chem. 9:623-632(1990).
RN   [3]
RP   SEQUENCE OF 38-85 AND 141-175.
RX   MEDLINE=85298214; PubMed=3862086;
RA   Gross J., Brauer A.W., Bringhurst R.F., Corbett C., Margolies M.N.;
RT   "An unusual bovine pancreatic protein exhibiting pH-dependent
RT   globule-fibril transformation and unique amino acid sequence.";
RL   Proc. Natl. Acad. Sci. U.S.A. 82:5627-5631(1985).
CC   -!- FUNCTION: Might act as an inhibitor of spontaneous calcium
CC       carbonate precipitation.
CC   -!- SUBUNIT: Cleaved to give an A chain and a B chain joined by a
CC       disulfide bond.
CC   -!- TISSUE SPECIFICITY: In pancreatic acinar cells.
CC   -!- SIMILARITY: Contains 1 C-type lectin family domain.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; M59794; AAA30750.1; -.
DR   PIR; A37194; A37194.
DR   HSSP; P05451; 1QDD.
DR   InterPro; IPR001304; Lectin_C.
DR   InterPro; IPR003990; Pancreatis_ac.
DR   Pfam; PF00059; lectin_c; 1.
DR   PRINTS; PR01504; PNCREATITSAP.
DR   SMART; SM00034; CLECT; 1.
DR   PROSITE; PS00615; C_TYPE_LECTIN_1; 1.
DR   PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
KW   Glycoprotein; Signal; Lectin.
FT   SIGNAL        1     26       POTENTIAL.
FT   PROPEP       27     37       POTENTIAL.
FT   CHAIN        38    175       LITHOSTATHINE.
FT   CHAIN        38    138       A CHAIN.
FT   CHAIN       141    175       B CHAIN.
FT   DOMAIN       38    173       C-TYPE LECTIN (LONG FORM).
FT   DISULFID     40     51       BY SIMILARITY.
FT   DISULFID     68    171       BY SIMILARITY.
FT   DISULFID    146    163       BY SIMILARITY.
FT   CONFLICT     84     85       EE -> FF (IN REF. 3).
SQ   SEQUENCE   175 AA;  19334 MW;  C270EE70B7E91D6A CRC64;

  Query Match             12.2%;  Score 136;  DB 1;  Length 175;
  Best Local Similarity   24.1%;  Pred. No. 3.2e-05;
  Matches   41;  Conservative   23;  Mismatches   64;  Indels   42;  Gaps    7;

Qy         5 LLSGQPVCRGGT---QRPCYKVIYFHDTSRRLNFEEAKEAC-RRDGGQLVSIESEDEQKL 60
             | | :  |  |:   :  || :       |     : :|  || :|  | |||: |  |:
Db        33 LPSARISCPSGSMAYRSHCYALFKTPKT-----WMDADIACQKRPSGHLVSVLSGAEESF 87

Qy        61 IEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEPSCGSEV 120
              :  : | | :  | ||||    :     | |    :  |     :   :  |  |
Db        88 VASLVRNNLNTQSDIWIGL---HDPTEGSEANAGGWEWISNDVLNYVAWETD-------- 136

Qy       121 CVVMYHQPSAPAGIGGP-----------YMFQWNDDRCNMKNNFICKYSD 159
                         || |  |               |: :|  ||:    ::||:|
Db       137 ----------PAAISSPGYCGSLSRSSGYL-KWRDHNCNLNLPYVCKFTD 175

RESULT 42
LIT1_MOUSE
ID   LIT1_MOUSE     STANDARD;      PRT;   165 AA.
AC   P43137;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Lithostathine 1 precursor (Pancreatic stone protein 1) (PSP)
DE   (Pancreatic thread protein 1) (PTP) (Islet of langerhans regenerating
DE   protein 1) (REG 1).
GN   REG1.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J;
RX   MEDLINE=93340209; PubMed=8340418;
RA   Unno M., Yonekura H., Nakagawara K.-I., Watanabe T., Miyashita H.,
RA   Moriizumi S., Okamoto H., Itoh T., Teraoka H.;
RT   "Structure, chromosomal localization, and expression of mouse reg
RT   genes, reg I and reg II. A novel type of reg gene, reg II, exists in
RT   the mouse genome.";
RL   J. Biol. Chem. 268:15974-15982(1993).
CC   -!- FUNCTION: Might act as an inhibitor of spontaneous calcium
CC       carbonate precipitation.
CC   -!- TISSUE SPECIFICITY: Expressed only in regenerating islets and
CC       normal exocrine pancreas, but not in normal pancreatic islets.
CC       Expressed strongly in pancreas, moderately in gallbladder, and
CC       weakly in liver.
CC   -!- SIMILARITY: Contains 1 C-type lectin family domain.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; D14010; BAA03111.1; -.
DR   PIR; A47148; A47148.
```

```
DR   HSSP; P05451; 1LIT.
DR   MGD; MGI:97895; Reg1.
DR   InterPro; IPR002353; AntifreezeII.
DR   InterPro; IPR001304; Lectin_C.
DR   InterPro; IPR003990; Pancreatis_ac.
DR   Pfam; PF00059; lectin_c; 1.
DR   PRINTS; PR00356; ANTIFREEZEII.
DR   PRINTS; PR01504; PNCREATITSAP.
DR   SMART; SM00034; CLECT; 1.
DR   PROSITE; PS00615; C_TYPE_LECTIN_1; 1.
DR   PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
KW   Glycoprotein; Signal; Lectin.
FT   SIGNAL        1     21       BY SIMILARITY.
FT   CHAIN        22    165       LITHOSTATHINE 1.
FT   DOMAIN       33    163       C-TYPE LECTIN (LONG FORM).
FT   DISULFID     35     46       BY SIMILARITY.
FT   DISULFID     63    161       BY SIMILARITY.
FT   DISULFID    136    153       BY SIMILARITY.
FT   CARBOHYD    129    129       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   165 AA;  18518 MW;  2950174AF5D666BA CRC64;

  Query Match             12.1%;  Score 135;  DB 1;  Length 165;
  Best Local Similarity   26.4%;  Pred. No. 3.6e-05;
  Matches   37;  Conservative   28;  Mismatches   51;  Indels   24;  Gaps    8;

Qy         20 CYKVIYFHDTSRRLNFEEAKEACRR-DGGQLVSIESEDEQKLIEKFIENLLPSDGDFWIG 78
              ||   ||  |  || : :|   |:  : | |||: |: |   :   |:    :| : | |
Db         46 CY---YF--TEDRLTWADADLFCQNMNSGYLVSVLSQAEGNFVASLIKESGTTDANVWTG 100

Qy         79 LRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEPSCGSE-VCVVMYHQPSAPAGIGGP 137
              |   : |::        : |: ||:  :::|    |:  :   ||       |  :  |
Db        101 L--HDPKRNRR------WHWSSGSLFLYKSWATGSPNSSNRGYCV------SLTSNTG-- 144

Qy        138 YMFQWNDDRCNMKNNFICKY 157
                :| || |: : :|:||:
Db        145 -YKKWKDDNCDAQYSFVCKF 163


RESULT 43
ACAL_ANSAN
ID   ACAL_ANSAN     STANDARD;      PRT;   132 AA.
AC   P83300;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Ansocalcin.
OS   Anser anser anser (Western graylag goose).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Anseriformes; Anatidae; Anser.
OX   NCBI_TaxID=8844;
RN   [1]
RP   SEQUENCE, FUNCTION, SUBUNIT, SUBCELLULAR LOCATION, AND MASS
RP   SPECTROMETRY.
RC   TISSUE=Eggshell matrix;
RX   MEDLINE=22439773; PubMed=12431998;
RA   Lakshminarayanan R., Valiyaveettil S., Rao V.S., Kini R.M.;
RT   "Purification, characterization, and in vitro mineralization studies
RT   of a novel goose eggshell matrix protein, ansocalcin.";
RL   J. Biol. Chem. 278:2928-2936(2003).
CC   -!- FUNCTION: Induces spherical aggregates of calcite crystals in
CC       vitro. Believed to play an active role in the eggshell
CC       calcification.
CC   -!- SUBUNIT: Homodimer or homotrimer.
CC   -!- SUBCELLULAR LOCATION: Eggshell matrix. May be present in the shell
CC       glands on the walls of oviduct and incorporated into the shell
CC       structure during its formation.
CC   -!- MASS SPECTROMETRY: MW=15342; METHOD=Electrospray.
CC   -!- SIMILARITY: Contains 1 C-type lectin family domain.
DR   InterPro; IPR002353; AntifreezeII.
DR   InterPro; IPR001304; Lectin_C.
DR   InterPro; IPR003990; Pancreatis_ac.
DR   Pfam; PF00059; lectin_c; 1.
DR   PRINTS; PR00356; ANTIFREEZEII.
DR   PRINTS; PR01504; PNCREATITSAP.
DR   SMART; SM00034; CLECT; 1.
DR   PROSITE; PS00615; C_TYPE_LECTIN_1; 1.
DR   PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
KW   Lectin.
FT   DOMAIN        1    132       C-TYPE LECTIN (LONG FORM) (BY
FT                                SIMILARITY).
FT   DISULFID      3     14       BY SIMILARITY.
FT   DISULFID     31    128       BY SIMILARITY.
FT   DISULFID    103    120       BY SIMILARITY.
SQ   SEQUENCE   132 AA;  15347 MW;  36CE42EA4572E6B7 CRC64;

  Query Match             12.1%;  Score 134.5;  DB 1;  Length 132;
  Best Local Similarity   26.6%;  Pred. No. 3.1e-05;
  Matches   38;  Conservative   21;  Mismatches   55;  Indels   29;  Gaps    7;

Qy         20 CYKVIYFHDTSRRLNFEEAKEACR--RDGGQLVSIESEDEQKLIEKFIENL--LPSDGDF 75
              ||   ||     : | : :|:  |:     |  | |: | :|   : :||       : :
Db         14 CYG--YF---GQELTWRKAEAWCKVIHAGCHLASLHSPEEHAAVARFIAKFQRREEEDNV 68

Qy         76 WIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEPSCGSEVCVVMYHQPSAPAGIG 135
              ||||    : :           | |||  ::  |  ||   | : | |:          |
Db         69 WIGLHHWNQAR----------VWIDGSKKRYSAWDDDELPRG-KYCTVLE---------G 108

Qy        136 GPYMFQWNDDRCNMKNNFICKYS 158
                    | |: |: :| |:||||
Db        109 SSGFMSWEDNACSERNPFVCKYS 131


RESULT 44
LEM1_PONPY
ID   LEM1_PONPY     STANDARD;      PRT;   372 AA.
AC   Q95235;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   L-selectin precursor (Lymph node homing receptor) (Leukocyte adhesion
DE   molecule-1) (LAM-1) (Leukocyte-endothelial cell adhesion molecule 1)
DE   (LECAM1) (CD62L).
GN   SELL.
OS   Pongo pygmaeus (Orangutan).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pongo.
OX   NCBI_TaxID=9600;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Budman J.I., Fu H., Johnson C.E., Thakur A.B., Berg E.L.,
RA   Tsurushita N.;
RL   Submitted (NOV-1996) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: Cell surface adhesion protein. Mediate the adherence of
CC       lymphocytes to endothelial cells of high endothelial venules in
CC       peripheral lymph nodes.
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC   -!- SIMILARITY: Belongs to the selectin/LECAM family.
CC   -!- SIMILARITY: Contains 1 C-type lectin family domain.
CC   -!- SIMILARITY: Contains 1 EGF-like domain.
CC   -!- SIMILARITY: Contains 2 Sushi (SCR) domains.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; U73729; AAB18247.1; -.
DR   HSSP; P14151; 1KJB.
DR   InterPro; IPR006209; EGF_like.
DR   InterPro; IPR006210; IEGF.
DR   InterPro; IPR001304; Lectin_C.
DR   InterPro; IPR002396; Selectin.
DR   InterPro; IPR000436; Sushi_SCR_CCP.
```

```
DR   Pfam; PF00008; EGF; 1.
DR   Pfam; PF00059; lectin_c; 1.
DR   Pfam; PF00084; sushi; 2.
DR   PRINTS; PR00343; SELECTIN.
DR   SMART; SM00032; CCP; 2.
DR   SMART; SM00034; CLECT; 1.
DR   SMART; SM00181; EGF; 1.
DR   PROSITE; PS00615; C_TYPE_LECTIN_1; 1.
DR   PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
DR   PROSITE; PS00022; EGF_1; 1.
DR   PROSITE; PS01186; EGF_2; 1.
DR   PROSITE; PS50026; EGF_3; 1.
KW   Cell adhesion; Transmembrane; Glycoprotein; EGF-like domain; Lectin;
KW   Selectin; Signal; Sushi; Repeat.
FT   SIGNAL        1     28       BY SIMILARITY.
FT   PROPEP       29     38       BY SIMILARITY.
FT   CHAIN        39    372       L-SELECTIN.
FT   DOMAIN       39    332       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    333    355       POTENTIAL.
FT   DOMAIN      356    372       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN       55    155       C-TYPE LECTIN (SHORT FORM).
FT   DOMAIN      156    192       EGF-LIKE.
FT   DOMAIN      196    255       SUSHI 1.
FT   DOMAIN      258    317       SUSHI 2.
FT   DISULFID     57    155       BY SIMILARITY.
FT   DISULFID    128    147       BY SIMILARITY.
FT   DISULFID    160    171       BY SIMILARITY.
FT   DISULFID    165    180       BY SIMILARITY.
FT   DISULFID    182    191       BY SIMILARITY.
FT   DISULFID    197    241       BY SIMILARITY.
FT   DISULFID    227    254       BY SIMILARITY.
FT   DISULFID    259    303       BY SIMILARITY.
FT   DISULFID    289    316       BY SIMILARITY.
FT   CARBOHYD     60     60       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    104    104       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    177    177       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    216    216       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    226    226       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    232    232       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    246    246       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    271    271       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   372 AA;  42118 MW;  6517DD22213FF15E CRC64;

  Query Match             12.1%;  Score 134.5;  DB 1;  Length 372;
  Best Local Similarity   27.4%;  Pred. No. 0.00011;
  Matches   43;  Conservative   33;  Mismatches   52;  Indels   29;  Gaps    9;

Qy          26 FHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLP-SDGDFWIGLRRREE 84
               :| : : :|:: |:  || :   ||:|::: |   || ::|  || |   :|||:|:
Db          41 YHYSEKPMNWQRARRFCRENYTDLVAIQNKAE---IE-YLEKTLPFSRSYYWIGIRK--- 93

Qy          85 KQSNSTACQDLYAW--TDGSIS-QFRNWYVDEPS--CGSEVCVVMYHQPSAPAGIGGPYM 139
                  :: |  |: |:: :  ||   ||:      | || :| : :  ||
Db          94 -------IGGIWTWVGTNKSLTEEAENWGDGEPNNKKNKEDCVEIYIKRNKDAG------ 140

Qy         140 FQWNDDRCNMKNNFICKYSDEKPAVPSREAEGEETEL 176
               :|||| |:     :|  :  :|   |    ||  |:
Db         141 -KWNDDACHKLKAALCYTASCQPW--SCSGHGECVEI 174


RESULT 45
LEM2_BOVIN
ID   LEM2_BOVIN     STANDARD;      PRT;   485 AA.
AC   P98107;
DT   01-FEB-1996 (Rel. 33, Created)
DT   01-FEB-1996 (Rel. 33, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   E-selectin precursor (Endothelial leukocyte adhesion molecule 1)
DE   (ELAM-1) (Leukocyte-endothelial cell adhesion molecule 2) (LECAM2).
GN   SELE.
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Adrenal gland;
RX   MEDLINE=93382537; PubMed=7690465;
RA   Nguyen M., Strubel N.A., Bischoff J.;
RT   "A role for sialyl Lewis-X/A glycoconjugates in capillary
RT   morphogenesis.";
RL   Nature 365:267-269(1993).
CC   -!- FUNCTION: Expressed on cytokine induced endothelial cells and
CC       mediates their binding to Leukocytes. The ligand recognized by
CC       ELAM-1 is sialyl-lewis X (alpha(1->3)fucosylated derivatives of
CC       polylactosamine that are found at the nonreducing termini of
CC       glycolipids).
CC   -!- FUNCTION: Involved in capillary morphogenesis by binding to a
CC       bovine capillary endothelial (BCE) cell sialyl Lewis-X and/or
CC       sialyl Lewis A-containing ligand (probably).
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC   -!- SIMILARITY: Belongs to the selectin/LECAM family.
CC   -!- SIMILARITY: Contains 1 C-type lectin family domain.
CC   -!- SIMILARITY: Contains 1 EGF-like domain.
CC   -!- SIMILARITY: Contains 4 Sushi (SCR) domains.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; L12039; AAA02991.1; -.
DR   PIR; S36772; S36772.
DR   HSSP; P16581; 1ESL.
DR   InterPro; IPR000742; EGF_2.
DR   InterPro; IPR006209; EGF_like.
DR   InterPro; IPR006210; IEGF.
DR   InterPro; IPR001304; Lectin_C.
DR   InterPro; IPR002396; Selectin.
DR   InterPro; IPR000436; Sushi_SCR_CCP.
DR   Pfam; PF00008; EGF; 1.
DR   Pfam; PF00059; lectin_c; 1.
DR   Pfam; PF00084; sushi; 4.
DR   PRINTS; PR00343; SELECTIN.
DR   SMART; SM00032; CCP; 4.
DR   SMART; SM00034; CLECT; 1.
DR   SMART; SM00181; EGF; 1.
DR   PROSITE; PS00615; C_TYPE_LECTIN_1; 1.
DR   PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
DR   PROSITE; PS00022; EGF_1; 1.
DR   PROSITE; PS01186; EGF_2; 1.
DR   PROSITE; PS50026; EGF_3; 1.
KW   Cell adhesion; Transmembrane; Glycoprotein; EGF-like domain; Lectin;
KW   Selectin; Signal; Sushi; Repeat.
FT   SIGNAL        1     22       BY SIMILARITY.
FT   CHAIN        23    485       E-SELECTIN.
FT   DOMAIN       23    430       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    431    453       POTENTIAL.
FT   DOMAIN      454    485       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN       23    140       C-TYPE LECTIN.
FT   DOMAIN      141    176       EGF-LIKE.
FT   DOMAIN      180    238       SUSHI 1.
FT   DOMAIN      241    300       SUSHI 2.
FT   DOMAIN      303    363       SUSHI 3.
FT   DOMAIN      366    422       SUSHI 4.
FT   DISULFID     41    139       BY SIMILARITY.
FT   DISULFID    112    131       BY SIMILARITY.
FT   DISULFID    144    155       BY SIMILARITY.
FT   DISULFID    149    164       BY SIMILARITY.
FT   DISULFID    166    175       BY SIMILARITY.
FT   DISULFID    181    224       BY SIMILARITY.
FT   DISULFID    210    237       BY SIMILARITY.
```

```
FT   DISULFID    242    286       BY SIMILARITY.
FT   DISULFID    272    299       BY SIMILARITY.
FT   DISULFID    304    349       BY SIMILARITY.
FT   DISULFID    335    362       BY SIMILARITY.
FT   DISULFID    367    408       BY SIMILARITY.
FT   DISULFID    394    421       BY SIMILARITY.
FT   CARBOHYD     61     61       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     79     79       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     88     88       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    161    161       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    203    203       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    265    265       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    312    312       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    316    316       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    379    379       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    401    401       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   485 AA;  53200 MW;  AE931C9B521E3904 CRC64;

  Query Match             12.1%;  Score 134.5;  DB 1;  Length 485;
  Best Local Similarity   26.0%;  Pred. No. 0.00015;
  Matches   40;  Conservative   31;  Mismatches   58;  Indels   25;  Gaps    7;

Qy         26 FHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFWIGLRRREEK 85
              :|  ::   :  ||||:: |::     ||:|::::|   ||         |   :|||:|:
Db         25 YHASTEMMTFEEARDYCQKTYTALVAIQNQEE---IEYLNSTFSYSPSYYWIGIRK---- 77

Qy         86 QSNSTACQDLYAW--TDGSIS-QFRNWYVDEPS--CGSEVCVVMYHQPSAPAGIGGPYMF 140
                | |      : |  |: |:: :  ||    ||:      | || :| :    :|
Db         78 -INGT-----WTWIGTNKSLTKEATNWAPGEPNNKQSDEDCVEIYIKREKDSG------- 124

Qy        141 QWNDDRCNMKNNFICKYSDEKPAVPSREAEGEET 174
              :|||::|  : . :|  :.  |       |  ||
Db        125 KWNDEKCTKQKLALCYKAACNPTPCGSHGECVET 158


RESULT 46
TETN_CARSP
ID   TETN_CARSP     STANDARD;      PRT;   166 AA.
AC   P26258;
DT   01-MAY-1992 (Rel. 22, Created)
DT   01-MAY-1992 (Rel. 22, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   Tetranectin-like protein.
OS   Carcharhinus springeri (Reef shark).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;
OC   Elasmobranchii; Galeomorphii; Galeoidea; Carcharhiniformes;
OC   Carcharhinidae; Carcharhinus.
OX   NCBI_TaxID=7809;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Neural arch cartilage;
RX   MEDLINE=93284081; PubMed=1304877;
RA   Neame P.J., Young C.N., Treep J.T.;
RT   "Primary structure of a protein isolated from reef shark
RT   (Carcharhinus springeri) cartilage that is similar to the mammalian
RT   C-type lectin homolog, tetranectin.";
RL   Protein Sci. 1:161-168(1992).
CC   -!- SIMILARITY: Contains 1 C-type lectin family domain.
DR   PIR; A37289; A37289.
DR   HSSP; P05452; 1HTN.
DR   InterPro; IPR001304; Lectin_C.
DR   Pfam; PF00059; lectin_c; 1.
DR   SMART; SM00034; CLECT; 1.
DR   PROSITE; PS00615; C_TYPE_LECTIN_1; 1.
DR   PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
KW   Lectin.
FT   DOMAIN       43    161       C-TYPE LECTIN.
FT   DISULFID     37     47
FT   DISULFID     64    160
FT   DISULFID    136    152
SQ   SEQUENCE   166 AA;  18432 MW;  53EF812DEA5C6119 CRC64;

  Query Match             12.0%;  Score 134;  DB 1;  Length 166;
  Best Local Similarity   26.3%;  Pred. No. 4.5e-05;
  Matches   40;  Conservative   23;  Mismatches   63;  Indels   26;  Gaps    6;

Qy          9 QPVCRGGTQ--RPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIE 66
              | ||  ||:  : ||        :    ::  | | |   || |    | ||   :  : :
Db         34 QTVCLKGTKIHKKCYLA-----SRGSKSYHAANEDCIAQGGTLSIPRSSDEGNSLRSYAK 88

Qy         67 NLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEPSCGS-EVCVVMY 125
                |.   |||||:        |    :  :   :|    : ||   :|  |: | ||
Db         89 KSLVGARDFWIGV--------NDMTTEGKFVDVNGLPITYFNWDRSKPVGGTRENCVA-- 138

Qy        126 HQPSAPAGIGGPYMFQWNDDRCNMKNNFICKY 157
                :: :| |     :|:|| |  :  :||:|
Db        139 ---ASTSGQG-----KWSDDVCRSEKRYICEY 162


RESULT 47
ECHA_ECHCA
ID   ECHA_ECHCA     STANDARD;      PRT;   133 AA.
AC   P81017;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Echicetin alpha subunit.
OS   Echis carinatus (Saw-scaled viper).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
OC   Viperidae; Viperinae; Echis.
OX   NCBI_TaxID=40353;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Venom;
RX   MEDLINE=97250657; PubMed=9163349;
RA   Polgar J., Magnenat E.M., Peitsch M.C., Wells T.N.C., Saqi M.S.A.,
RA   Clemetson K.J.;
RT   "Amino acid sequence of the alpha subunit and computer modelling of
RT   the alpha and beta subunits of echicetin from the venom of Echis
RT   carinatus (saw-scaled viper).";
RL   Biochem. J. 323:533-537(1997).
RN   [2]
RP   CHARACTERIZATION.
RC   TISSUE=Venom;
RX   MEDLINE=93244424; PubMed=8481512;
RA   Peng M., Lu W., Beviglia L., Niewiarowski S., Kirby E.P.;
RT   "Echicetin: a snake venom protein that inhibits binding of von
RT   Willebrand factor and alboaggregins to platelet glycoprotein Ib.";
RL   Blood 81:2321-2328(1993).
CC   -!- FUNCTION: Binds to platelet GPIb and inhibits platelet
CC       agglutination.
CC   -!- SUBUNIT: Heterodimer of alpha and beta subunits.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- SIMILARITY: Contains 1 C-type lectin family domain.
DR   HSSP; P23806; 1IXX.
DR   InterPro; IPR001304; Lectin_C.
DR   InterPro; IPR003990; Pancreatis_ac.
DR   Pfam; PF00059; lectin_c; 1.
DR   PRINTS; PR01504; PNCREATITSAP.
DR   SMART; SM00034; CLECT; 1.
DR   PROSITE; PS00615; C_TYPE_LECTIN_1; FALSE_NEG.
DR   PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
KW   Lectin.
FT   DISULFID      4     15       BY SIMILARITY.
FT   DISULFID     31    127       BY SIMILARITY.
FT   DISULFID    102    119       BY SIMILARITY.
FT   DISULFID     81     81       INTERCHAIN (WITH C-75 IN BETA CHAIN) (BY
FT                                SIMILARITY).
SQ   SEQUENCE   133 AA;  15803 MW;  8F11C2D0BDC70B16 CRC64;

  Query Match             12.0%;  Score 133.5;  DB 1;  Length 133;
  Best Local Similarity   29.1%;  Pred. No. 3.8e-05;
  Matches   41;  Conservative   21;  Mismatches   48;  Indels   31;  Gaps    8;

Qy         20 CYKVIYFHDTSRRLNFEEAKEACRR-DGGQLVSIESEDEQKLIEKFIENLLPS---DGDF 75
```

```
                ||::       |   ::||::  | : ||| ||||||  : : : : |    ||     :
Db          15 CYQLF------RLKTWDEAEKYCNQWDGGHLVSIESNAKAEFVAQLISRKLPKSAIEDRV 68

Qy          76 WIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEPSCGSEVCVVMYHQPSAPAGIG 135
               ||||| | :::     |  |  ||| |   :   :|  |:     | |:  |
Db          69 WIGLRDRSKREQ----CGHL--WTDNSFVHYE--HVVPPT----KCFVLERQTE------ 110

Qy         136 GPYMFQWNDDRCNMKNNFICK 156
                    :|    |  |  |:||
Db         111 ---FRKWIAVNCEFKFPFVCK 128


RESULT 48
LECI_HUMAN
ID   LECI_HUMAN     STANDARD;      PRT;   311 AA.
AC   P07307; O00448; Q03969;
DT   01-APR-1988 (Rel. 07, Created)
DT   01-APR-1988 (Rel. 07, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   Asialoglycoprotein receptor 2 (Hepatic lectin H2) (ASGP-R) (ASGPR).
GN   ASGR2.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORM 1).
RX   MEDLINE=86016723; PubMed=3863106;
RA   Spiess M., Lodish H.F.;
RT   "Sequence of a second human asialoglycoprotein receptor: conservation
RT   of two receptor genes during evolution.";
RL   Proc. Natl. Acad. Sci. U.S.A. 82:6465-6469(1985).
RN   [2]
RP   SEQUENCE FROM N.A. (ISOFORMS 2 AND 3).
RC   TISSUE=Liver;
RX   MEDLINE=92184202; PubMed=1371982;
RA   Paietta E., Stockert R.J., Racevskis J.;
RT   "Differences in the abundance of variably spliced transcripts for the
RT   second asialoglycoprotein receptor polypeptide, H2, in normal and
RT   transformed human liver.";
RL   Hepatology 15:395-402(1992).
RN   [3]
RP   SEQUENCE FROM N.A. (ISOFORM 3).
RC   TISSUE=Liver;
RX   MEDLINE=22388257; PubMed=12477932;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length
RT   human and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC   -!- FUNCTION: Mediates the endocytosis of plasma glycoproteins to
CC       which the terminal sialic acid residue on their complex
CC       carbohydrate moieties has been removed. The receptor recognizes
CC       terminal galactose and N-acetylgalactosamine units. After ligand
CC       binding to the receptor, the resulting complex is internalized and
CC       transported to a sorting organelle, where receptor and ligand are
CC       disassociated. The receptor then returns to the cell membrane
CC       surface.
CC   -!- SUBUNIT: The functioning ligand-binding unit of this receptor is
CC       thought to be at least a dimer.
CC   -!- SUBCELLULAR LOCATION: Type II membrane protein.
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=3;
CC       Name=1;
CC         IsoId=P07307-1; Sequence=Displayed;
CC       Name=2;
CC         IsoId=P07307-2; Sequence=VSP_003060;
CC       Name=3;
CC         IsoId=P07307-3; Sequence=VSP_003060, VSP_003061;
CC   -!- TISSUE SPECIFICITY: Expressed exclusively in hepatic parenchymal
CC       cells.
CC   -!- MISCELLANEOUS: Calcium is required for ligand binding.
CC   -!- SIMILARITY: Contains 1 C-type lectin family domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M11025; AAB59519.1; -.
DR   EMBL; U97197; AAB58308.1; -.
DR   EMBL; X55283; CAA38997.1; -.
DR   EMBL; BC017251; AAH17251.1; -.
DR   PIR; A25179; LNHU2A.
DR   HSSP; P06734; 1HLI.
DR   Genew; HGNC:743; ASGR2.
DR   MIM; 108361; -.
DR   GO; GO:0004873; F:asialoglycoprotein receptor activity; TAS.
DR   GO; GO:0007166; P:cell surface receptor linked signal transdu. . .; TAS.
DR   InterPro; IPR002353; AntifreezeII.
DR   InterPro; IPR001304; Lectin_C.
DR   InterPro; IPR005640; lectin_N.
DR   Pfam; PF00059; lectin_c; 1.
DR   Pfam; PF03954; lectin_N; 1.
DR   PRINTS; PR00356; ANTIFREEZEII.
DR   SMART; SM00034; CLECT; 1.
DR   PROSITE; PS00615; C_TYPE_LECTIN_1; 1.
DR   PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
KW   Lectin; Glycoprotein; Receptor; Endocytosis; Transmembrane;
KW   Calcium; Signal-anchor; Phosphorylation; Alternative splicing.
FT   DOMAIN        1     58       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     59     79       SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
FT                                (POTENTIAL).
FT   DOMAIN       80    311       EXTRACELLULAR (POTENTIAL).
FT   DOMAIN      176    302       C-TYPE LECTIN (LONG FORM).
FT   SITE          5      8       ENDOCYTOSIS SIGNAL (POTENTIAL).
FT   DISULFID    177    188       BY SIMILARITY.
FT   DISULFID    205    300       BY SIMILARITY.
FT   DISULFID    278    292       BY SIMILARITY.
FT   CARBOHYD    102    102       N-LINKED (GLCNAC. . .).
FT   CARBOHYD    170    170       N-LINKED (GLCNAC. . .).
FT   CARBOHYD    305    305       N-LINKED (GLCNAC. . .).
FT   MOD_RES      12     12       PHOSPHORYLATION.
FT   VARSPLIC     24     42       Missing (in isoform 2 and isoform 3).
FT                                /FTId=VSP_003060.
FT   VARSPLIC     82     86       Missing (in isoform 3).
FT                                /FTId=VSP_003061.
SQ   SEQUENCE   311 AA;  35191 MW;  82C78FEC8FEBA316 CRC64;

  Query Match             12.0%;  Score 133.5;  DB 1;  Length 311;
  Best Local Similarity   27.4%;  Pred. No. 0.00011;
  Matches   48;  Conservative   23;  Mismatches   51;  Indels   53;  Gaps   10;

Qy          15 GTQRPCYKV--------IYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIE 66
               |:|| | |          |:   | :  : ||:: |: :   || | | :||| |   ::
Db         171 GSQRTCCPVNWVEHQGSCYWFSHSGKA-WAEAEKYCQLENAHLVVINSWEEQKFI---VQ 226

Qy          67 NLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGS--ISQFRNWYVDEPS-------CG 117
```

```
                :  |   : ||||          |    : | ||:     ::|| | :|         |
Db        227 HTNPF--NTWIGL----------TDSDGSWKWVDGTDYRHNYKNWAVTQPDNWHGHELGG 274

Qy        118 SEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGE 172
                || ||  :  ||             :|||| |     ::|:          | | ||
Db        275 SEDCVEV--QPDG----------RWNDDFCLQVYRWVCE--------KRRNATGE 309


RESULT 49
LEM1_HUMAN
ID   LEM1_HUMAN     STANDARD;      PRT;   372 AA.
AC   P14151; P15023;
DT   01-JAN-1990 (Rel. 13, Created)
DT   01-FEB-1991 (Rel. 17, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   L-selectin precursor (Lymph node homing receptor) (Leukocyte adhesion
DE   molecule-1) (LAM-1) (Leukocyte surface antigen Leu-8) (TQ1) (gp90-MEL)
DE   (Leukocyte-endothelial cell adhesion molecule 1) (LECAM1) (CD62L).
GN   SELL OR LYAM1 OR LNHR.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=89315837; PubMed=2664786;
RA   Siegelman M.H., Weissman I.L.;
RT   "Human homologue of mouse lymph node homing receptor: evolutionary
RT   conservation at tandem cell interaction domains.";
RL   Proc. Natl. Acad. Sci. U.S.A. 86:5562-5566(1989).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Tonsil;
RX   MEDLINE=89310350; PubMed=2473156;
RA   Tedder T.F., Isaacs C.M., Ernst T.J., Demetri G.D., Adler D.A.,
RA   Disteche C.M.;
RT   "Isolation and chromosomal localization of cDNAs encoding a novel
RT   human lymphocyte cell surface molecule, LAM-1. Homology with the
RT   mouse lymphocyte homing receptor and other human adhesion proteins.";
RL   J. Exp. Med. 170:123-133(1989).
RN   [3]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=90044046; PubMed=2509939;
RA   Camerini D., James S.P., Stamenkovic I., Seed B.;
RT   "Leu-8/TQ1 is the human equivalent of the Mel-14 lymph node homing
RT   receptor.";
RL   Nature 342:78-82(1989).
RN   [4]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=89308881; PubMed=2663882;
RA   Bowen B.R., Nguyen T., Lasky L.A.;
RT   "Characterization of a human homologue of the murine peripheral lymph
RT   node homing receptor.";
RL   J. Cell Biol. 109:421-427(1989).
RN   [5]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=90243637; PubMed=1692315;
RA   Ord D.C., Ernst T.J., Zhou L.J., Rambaldi A., Spertini O.,
RA   Griffin J., Tedder T.F.;
RT   "Structure of the gene encoding the human leukocyte adhesion
RT   molecule-1 (TQ1, Leu-8) of lymphocytes and neutrophils.";
RL   J. Biol. Chem. 265:7760-7767(1990).
RN   [6]
RP   SEQUENCE FROM N.A.
RA   Rieder M.J., Carrington D.P., da Ponte S.H., Hastings N.C.,
RA   Ahearn M.O., Kuldanek S.A., Rajkumar N., Toth E.J., Yi Q.,
RA   Nickerson D.A.;
RL   Submitted (FEB-2003) to the EMBL/GenBank/DDBJ databases.
RN   [7]
RP   3D-STRUCTURE MODELING.
RX   MEDLINE=96074584; PubMed=7488174;
RA   Bajorath J., Aruffo A.;
RT   "A template for generation and comparison of three-dimensional
RT   selectin models.";
RL   Biochem. Biophys. Res. Commun. 216:1018-1023(1995).
CC   -!- FUNCTION: Cell surface adhesion protein. Mediate the adherence of
CC       lymphocytes to endothelial cells of high endothelial venules in
CC       peripheral lymph nodes.
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC   -!- SIMILARITY: Belongs to the selectin/LECAM family.
CC   -!- SIMILARITY: Contains 1 C-type lectin family domain.
CC   -!- SIMILARITY: Contains 1 EGF-like domain.
CC   -!- SIMILARITY: Contains 2 Sushi (SCR) domains.
CC   -!- DATABASE: NAME=PROW; NOTE=CD guide CD62L entry;
CC       WWW="http://www.ncbi.nlm.nih.gov/prow/cd/cd621.htm".
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M25280; AAC63053.1; -.
DR   EMBL; X16150; CAA34275.1; -.
DR   EMBL; X17519; CAB43536.1; -.
DR   EMBL; X17519; CAB43537.1; ALT_SEQ.
DR   EMBL; X16070; CAA34203.1; ALT_SEQ.
DR   EMBL; M32414; AAB60700.1; -.
DR   EMBL; M32406; AAB60700.1; JOINED.
DR   EMBL; M32407; AAB60700.1; JOINED.
DR   EMBL; M32408; AAB60700.1; JOINED.
DR   EMBL; M32409; AAB60700.1; JOINED.
DR   EMBL; M32410; AAB60700.1; JOINED.
DR   EMBL; M32411; AAB60700.1; JOINED.
DR   EMBL; M32412; AAB60700.1; JOINED.
DR   EMBL; M32413; AAB60700.1; JOINED.
DR   EMBL; AY233976; AAO48272.1; -.
DR   PDB; 1KJB; 03-APR-96.
DR   GlycoSuiteDB; P14151; -.
DR   Genew; HGNC:10720; SELL.
DR   MIM; 153240; -.
DR   GO; GO:0005887; C:integral to plasma membrane; TAS.
DR   GO; GO:0008337; F:selectin; TAS.
DR   GO; GO:0007155; P:cell adhesion; TAS.
DR   InterPro; IPR006209; EGF_like.
DR   InterPro; IPR006210; IEGF.
DR   InterPro; IPR001304; Lectin_C.
DR   InterPro; IPR002396; Selectin.
DR   InterPro; IPR000436; Sushi_SCR_CCP.
DR   Pfam; PF00008; EGF; 1.
DR   Pfam; PF00059; lectin_c; 1.
DR   Pfam; PF00084; sushi; 2.
DR   PRINTS; PR00343; SELECTIN.
DR   SMART; SM00032; CCP; 2.
DR   SMART; SM00034; CLECT; 1.
DR   SMART; SM00181; EGF; 1.
DR   PROSITE; PS00615; C_TYPE_LECTIN_1; 1.
DR   PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
DR   PROSITE; PS00022; EGF_1; 1.
DR   PROSITE; PS01186; EGF_2; 1.
DR   PROSITE; PS50026; EGF_3; 1.
KW   Cell adhesion; Transmembrane; Glycoprotein; EGF-like domain; Lectin;
KW   Selectin; Signal; Sushi; Repeat; 3D-structure.
FT   SIGNAL       1     28
FT   PROPEP      29     38
FT   CHAIN       39    372       L-SELECTIN.
FT   DOMAIN      39    332       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   333    355       POTENTIAL.
FT   DOMAIN     356    372       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN      55    155       C-TYPE LECTIN (SHORT FORM).
FT   DOMAIN     156    192       EGF-LIKE.
FT   DOMAIN     196    255       SUSHI 1.
FT   DOMAIN     258    317       SUSHI 2.
FT   DISULFID    57    155       BY SIMILARITY.
```

```
FT   DISULFID    128    147       BY SIMILARITY.
FT   DISULFID    160    171       BY SIMILARITY.
FT   DISULFID    165    180       BY SIMILARITY.
FT   DISULFID    182    191       BY SIMILARITY.
FT   DISULFID    197    241       BY SIMILARITY.
FT   DISULFID    227    254       BY SIMILARITY.
FT   DISULFID    259    303       BY SIMILARITY.
FT   DISULFID    289    316       BY SIMILARITY.
FT   CARBOHYD     60     60       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    104    104       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    177    177       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    232    232       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    246    246       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    271    271       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    311    311       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CONFLICT     37     37       D -> Y (IN REF. 4).
FT   CONFLICT    178    178       Y -> H (IN REF. 4).
FT   CONFLICT    193    193       F -> L (IN REF. 1 AND 4).
FT   CONFLICT    213    213       P -> S (IN REF. 3).
FT   CONFLICT    214    214       L -> F (IN REF. 4).
FT   CONFLICT    218    220       SFS -> NFN (IN REF. 2).
FT   CONFLICT    242    242       G -> E (IN REF. 2).
FT   STRAND       40     43
FT   STRAND       48     48
FT   HELIX        50     60
FT   STRAND       63     64
FT   HELIX        70     79
FT   STRAND       87     89
FT   STRAND       91     94
FT   TURN         95     96
FT   STRAND       97    100
FT   TURN        101    103
FT   STRAND      106    106
FT   TURN        109    111
FT   STRAND      114    114
FT   TURN        116    117
FT   TURN        124    125
FT   STRAND      128    131
FT   TURN        133    134
FT   TURN        139    140
FT   STRAND      142    145
FT   TURN        147    148
FT   STRAND      151    157
SQ   SEQUENCE   372 AA;  42187 MW;  6EA9918ECA2D3643 CRC64;

  Query Match             11.9%;  Score 132.5;  DB 1;  Length 372;
  Best Local Similarity   27.4%;  Pred. No. 0.00017;
  Matches   43;  Conservative   33;  Mismatches   52;  Indels   29;  Gaps    9;

Qy         26 FHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLP-SDGDFWIGLRRREE 84
              :| : : :|:: |:  || :  ||:|::: |   || ::|  || |   :|||:|:
Db         41 YHYSEKPMNWQRARRFCRDNYTDLVAIQNKAE---IE-YLEKTLPFSRSYYWIGIRK--- 93

Qy         85 KQSNSTACQDLYAW--TDGSIS-QFRNWYVDEPS--CGSEVCVVMYHQPSAPAGIGGPYM 139
                        :: |  |: |:: :  ||   ||:     | || :| : :  ||
Db         94 -------IGGIWTWVGTNKSLTEEAENWGDGEPNNKKNKEDCVEIYIKRNKDAG------ 140

Qy        140 FQWNDDRCNMKNNFICKYSDEKPAVPSREAEGEETEL 176
              :|||| |:     :|  :  :|   |    ||  |:
Db        141 -KWNDDACHKLKAALCYTASCQPW--SCSGHGECVEI 174


RESULT 50
LEM1_PANTR
ID   LEM1_PANTR     STANDARD;      PRT;   372 AA.
AC   Q95237;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   L-selectin precursor (Lymph node homing receptor) (Leukocyte adhesion
DE   molecule-1) (LAM-1) (Leukocyte-endothelial cell adhesion molecule 1)
DE   (LECAM1) (CD62L).
GN   SELL.
OS   Pan troglodytes (Chimpanzee).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
OX   NCBI_TaxID=9598;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Budman J.I., Fu H., Johnson C.E., Thakur A.B., Berg E.L.,
RA   Tsurushita N.;
RT   "Cloning of the cDNA encoding L-selectin from nonhuman primates.";
RL   Submitted (NOV-1996) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: Cell surface adhesion protein. Mediate the adherence
CC       of lymphocytes to endothelial cells of high endothelial
CC       venules in peripheral lymph nodes.
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC   -!- SIMILARITY: Belongs to the selectin/LECAM family.
CC   -!- SIMILARITY: Contains 1 C-type lectin family domain.
CC   -!- SIMILARITY: Contains 1 EGF-like domain.
CC   -!- SIMILARITY: Contains 2 Sushi (SCR) domains.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; U73728; AAB18248.1; -.
DR   HSSP; P14151; 1KJB.
DR   InterPro; IPR006209; EGF_like.
DR   InterPro; IPR006210; IEGF.
DR   InterPro; IPR001304; Lectin_C.
DR   InterPro; IPR002396; Selectin.
DR   InterPro; IPR000436; Sushi_SCR_CCP.
DR   Pfam; PF00008; EGF; 1.
DR   Pfam; PF00059; lectin_c; 1.
DR   Pfam; PF00084; sushi; 2.
DR   PRINTS; PR00343; SELECTIN.
DR   SMART; SM00032; CCP; 2.
DR   SMART; SM00034; CLECT; 1.
DR   SMART; SM00181; EGF; 1.
DR   PROSITE; PS00615; C_TYPE_LECTIN_1; 1.
DR   PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
DR   PROSITE; PS00022; EGF_1; 1.
DR   PROSITE; PS01186; EGF_2; 1.
DR   PROSITE; PS50026; EGF_3; 1.
KW   Cell adhesion; Transmembrane; Glycoprotein; EGF-like domain; Lectin;
KW   Selectin; Signal; Sushi; Repeat.
FT   SIGNAL       1     28       BY SIMILARITY.
FT   PROPEP      29     38       BY SIMILARITY.
FT   CHAIN       39    372       L-SELECTIN.
FT   DOMAIN      39    332       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   333    355       POTENTIAL.
FT   DOMAIN     356    372       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN      55    155       C-TYPE LECTIN (SHORT FORM).
FT   DOMAIN     156    192       EGF-LIKE.
FT   DOMAIN     196    255       SUSHI 1.
FT   DOMAIN     258    317       SUSHI 2.
FT   DISULFID    57    155       BY SIMILARITY.
FT   DISULFID   128    147       BY SIMILARITY.
FT   DISULFID   160    171       BY SIMILARITY.
FT   DISULFID   165    180       BY SIMILARITY.
FT   DISULFID   182    191       BY SIMILARITY.
FT   DISULFID   197    241       BY SIMILARITY.
FT   DISULFID   227    254       BY SIMILARITY.
FT   DISULFID   259    303       BY SIMILARITY.
FT   DISULFID   289    316       BY SIMILARITY.
FT   CARBOHYD    60     60       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   104    104       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   177    177       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   216    216       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   232    232       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   246    246       N-LINKED (GLCNAC. . .) (POTENTIAL).
```

```
FT   CARBOHYD    271    271       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   372 AA;  42188 MW;  6EA991802A2D3643 CRC64;

  Query Match             11.9%;  Score 132.5;  DB 1;  Length 372;
  Best Local Similarity   27.4%;  Pred. No. 0.00017;
  Matches   43;  Conservative   33;  Mismatches   52;  Indels   29;  Gaps    9;

Qy         26 FHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLP-SDGDFWIGLRRREE 84
              :| : : :|:: |:  || :   ||:|::: |   || ::|  || |   :|||:|:
Db         41 YHYSEKPMNWQRARRFCRDNYTDLVAIQNKAE---IE-YLEKTLPFSRSYYWIGIRK--- 93

Qy         85 KQSNSTACQDLYAW--TDGSIS-QFRNWYVDEPS--CGSEVCVVMYHQPSAPAGIGGPYM 139
                        :: |  |: |:: :  ||   ||:     | || :| : :  ||
Db         94 -------IGGIWTWVGTNKSLTEEAENWGDGEPNNKKNKEDCVEIYIKRNKDAG------ 140

Qy        140 FQWNDDRCNMKNNFICKYSDEKPAVPSREAEGEETEL 176
               :|||| |:     :|  :  :|   |    ||  |:
Db        141 -KWNDDACHKLKAALCYTASCQPW--SCSGHGECVEI 174


Search completed: September  9, 2004, 22:36:48
Job time : 25 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: September 9, 2004, 22:36:28 ; Search time 32 Seconds
(without alignments)
332.342 Million cell updates/sec

Title: US-09-887-855-2_COPY_22_227
Perfect score: 1115
Sequence: 1 ATGRLLSGQPVCRGGTQRPC.....EEDAKKTFKESREAALNLAY 206

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 50 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1105	99.1	374	4	US-09-489-847-166	Sequence 166, App
2	1101	98.7	382	4	US-09-907-794A-137	Sequence 137, App
3	1101	98.7	382	4	US-09-905-125A-137	Sequence 137, App
4	1101	98.7	382	4	US-09-902-775A-137	Sequence 137, App
5	939.5	84.3	260	4	US-09-638-203-3	Sequence 3, Appli
6	561.5	50.4	273	4	US-09-638-203-2	Sequence 2, Appli
7	240	21.5	81	4	US-09-489-847-325	Sequence 325, App
8	240	21.5	82	4	US-09-489-847-234	Sequence 234, App
9	240	21.5	115	4	US-09-489-847-324	Sequence 324, App
10	182	16.3	1456	4	US-09-976-594-168	Sequence 168, App
11	178.5	16.0	1455	3	US-08-840-062-5	Sequence 5, Appli
12	176.5	15.8	1479	3	US-08-840-062-4	Sequence 4, Appli
13	174	15.6	1257	1	US-08-340-428B-49	Sequence 49, Appl
14	174	15.6	2409	6	5180808-2	Patent No. 5180808
15	170.5	15.3	1479	3	US-08-840-062-2	Sequence 2, Appli
16	152	13.6	912	5	PCT-US95-03747-2	Sequence 2, Appli
17	151	13.5	197	4	US-09-602-877A-99	Sequence 99, Appl
18	149.5	13.4	455	4	US-09-866-028-50	Sequence 50, Appl
19	149	13.4	174	1	US-07-641-971B-1	Sequence 1, Appli
20	149	13.4	174	1	US-07-781-248A-1	Sequence 1, Appli
21	149	13.4	320	1	US-08-365-103B-10	Sequence 10, Appl
22	149	13.4	321	1	US-08-365-103B-8	Sequence 8, Appli
23	148.5	13.3	652	2	US-08-751-305-2	Sequence 2, Appli
24	144.5	13.0	404	4	US-09-517-605-2	Sequence 2, Appli
25	143.5	12.9	372	2	US-08-513-278-4	Sequence 4, Appli
26	143.5	12.9	372	6	5514582-4	Patent No. 5514582
27	142	12.7	1487	3	US-08-840-062-7	Sequence 7, Appli
28	141.5	12.7	125	3	US-08-722-126A-7	Sequence 7, Appli
29	141.5	12.7	125	5	PCT-US95-04258-7	Sequence 7, Appli
30	141.5	12.7	287	1	US-08-365-103B-4	Sequence 4, Appli
31	141.5	12.7	300	1	US-08-365-103B-6	Sequence 6, Appli
32	141.5	12.7	327	1	US-08-365-103B-2	Sequence 2, Appli
33	138.5	12.4	117	6	5514582-7	Patent No. 5514582
34	138.5	12.4	119	1	US-08-340-539A-12	Sequence 12, Appl
35	137.5	12.3	110	6	5514582-12	Patent No. 5514582
36	137.5	12.3	123	6	5514582-19	Patent No. 5514582
37	137.5	12.3	492	4	US-09-724-864-39	Sequence 39, Appl
38	135	12.1	238	3	US-09-111-470-8	Sequence 8, Appli
39	134.5	12.1	128	4	US-09-535-521-8	Sequence 8, Appli
40	134.5	12.1	139	4	US-09-535-521-11	Sequence 11, Appl
41	134.5	12.1	141	4	US-09-535-521-14	Sequence 14, Appl
42	134.5	12.1	187	4	US-09-535-521-17	Sequence 17, Appl
43	134.5	12.1	208	4	US-09-535-521-20	Sequence 20, Appl
44	134.5	12.1	292	4	US-09-535-521-2	Sequence 2, Appli
45	134.5	12.1	292	4	US-09-535-521-5	Sequence 5, Appli
46	133.5	12.0	110	6	5514582-9	Patent No. 5514582
47	133.5	12.0	119	1	US-08-340-539A-13	Sequence 13, Appl
48	133.5	12.0	133	1	US-07-893-929A-9	Sequence 9, Appli
49	133.5	12.0	133	5	PCT-US92-10344-9	Sequence 9, Appli
50	133.5	12.0	287	3	US-09-111-470-6	Sequence 6, Appli

ALIGNMENTS

RESULT 1
US-09-489-847-166
; Sequence 166, Application US/09489847
; Patent No. 6476195
; GENERAL INFORMATION:
; APPLICANT: Rosen et al
; TITLE OF INVENTION: 98 Human Secreted Proteins
; FILE REFERENCE: PZ031P1
; CURRENT APPLICATION NUMBER: US/09/489,847
; CURRENT FILING DATE: 2000-01-24
; EARLIER APPLICATION NUMBER: PCT/US99/17130
; EARLIER FILING DATE: 1999-07-29
; EARLIER APPLICATION NUMBER: 60/094,657
; EARLIER FILING DATE: 1998-07-30
; EARLIER APPLICATION NUMBER: 60/095,486
; EARLIER FILING DATE: 1998-08-05
; EARLIER APPLICATION NUMBER: 60/096,319
; EARLIER FILING DATE: 1998-08-12
; EARLIER APPLICATION NUMBER: 60/095,454
; EARLIER FILING DATE: 1998-08-06
; EARLIER APPLICATION NUMBER: 60/095,455
; EARLIER FILING DATE: 1998-08-06
; NUMBER OF SEQ ID NOS: 376
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 166
; LENGTH: 374
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (84)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (112)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-489-847-166

Query Match 99.1%; Score 1105; DB 4; Length 374;
Best Local Similarity 99.0%; Pred. No. 1.9e-107;
Matches 204; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy          1 ATGRLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKL 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ATGRLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKL 81
```

```
Qy         61 IEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEPSCGSEV 120
              || |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db         82 IEXFIENLLPSDGDFWIGLRRREEKQSNSTXCQDLYAWTDGSISQFRNWYVDEPSCGSEV 141

Qy        121 CVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGEETELTTPV 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 CVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGEETELTTPV 201

Qy        181 LPEETQEEDAKKTFKESREAALNLAY 206
              |||||||||||||||||||||||||
Db        202 LPEETQEEDAKKTFKESREAALNLAY 227
```

RESULT 2
US-09-907-794A-137
; Sequence 137, Application US/09907794A
; Patent No. 6635468
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/907,794A
; CURRENT FILING DATE: 2001-07-17
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 137
; LENGTH: 382
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-907-794A-137

Query Match 98.7%; Score 1101; DB 4; Length 382;
Best Local Similarity 96.3%; Pred. No. 5.2e-107;
Matches 206; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

```
Qy          1 ATGRLLS--------GQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 52
              |||||||        ||||||||||||||||||||||||||||||||||||||||||||
Db         22 ATGRLLSASDLDLRGGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 81

Qy         53 ESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVD 112
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 ESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVD 141

Qy        113 EPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGE 172
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 EPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGE 201

Qy        173 ETELTTPVLPEETQEEDAKKTFKESREAALNLAY 206
              |||||||||||||||||||||||||||||||||
Db        202 ETELTTPVLPEETQEEDAKKTFKESREAALNLAY 235
```

RESULT 3
US-09-905-125A-137
; Sequence 137, Application US/09905125A
; Patent No. 6664376
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/905,125A
; CURRENT FILING DATE: 2001-07-12

```
;  PRIOR APPLICATION NUMBER: PCT/US00/04414
;  PRIOR FILING DATE: 2000-02-22
;  PRIOR APPLICATION NUMBER: US 60/143,048
;  PRIOR FILING DATE: 1999-07-07
;  PRIOR APPLICATION NUMBER: US 60/145,698
;  PRIOR FILING DATE: 1999-07-26
;  PRIOR APPLICATION NUMBER: US 60/146,222
;  PRIOR FILING DATE: 1999-07-28
;  PRIOR APPLICATION NUMBER: PCT/US99/20594
;  PRIOR FILING DATE: 1999-09-08
;  PRIOR APPLICATION NUMBER: PCT/US99/20944
;  PRIOR FILING DATE: 1999-09-13
;  PRIOR APPLICATION NUMBER: PCT/US99/21090
;  PRIOR FILING DATE: 1999-09-15
;  PRIOR APPLICATION NUMBER: PCT/US99/21547
;  PRIOR FILING DATE: 1999-09-15
;  PRIOR APPLICATION NUMBER: PCT/US99/23089
;  PRIOR FILING DATE: 1999-10-05
;  PRIOR APPLICATION NUMBER: PCT/US99/28214
;  PRIOR FILING DATE: 1999-11-29
;  PRIOR APPLICATION NUMBER: PCT/US99/28313
;  PRIOR FILING DATE: 1999-11-30
;  PRIOR APPLICATION NUMBER: PCT/US99/28564
;  PRIOR FILING DATE: 1999-12-02
;  PRIOR APPLICATION NUMBER: PCT/US99/28565
;  PRIOR FILING DATE: 1999-12-02
;  PRIOR APPLICATION NUMBER: PCT/US99/30095
;  PRIOR FILING DATE: 1999-12-16
;  PRIOR APPLICATION NUMBER: PCT/US99/30911
;  PRIOR FILING DATE: 1999-12-20
;  PRIOR APPLICATION NUMBER: PCT/US99/30999
;  PRIOR FILING DATE: 1999-12-20
;  PRIOR APPLICATION NUMBER: PCT/US00/00219
;  PRIOR FILING DATE: 2000-01-05
;  NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 137
;   LENGTH: 382
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-905-125A-137

  Query Match             98.7%;  Score 1101;  DB 4;  Length 382;
  Best Local Similarity   96.3%;  Pred. No. 5.2e-107;
  Matches  206;  Conservative    0;  Mismatches    0;  Indels    8;  Gaps    1;

Qy          1 ATGRLLS--------GQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI  52
              |||||||        ||||||||||||||||||||||||||||||||||||||||||||
Db         22 ATGRLLSASDLDLRGGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI  81

Qy         53 ESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVD  112
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 ESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVD  141

Qy        113 EPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGE  172
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 EPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGE  201

Qy        173 ETELTTPVLPEETQEEDAKKTFKESREAALNLAY  206
              |||||||||||||||||||||||||||||||||
Db        202 ETELTTPVLPEETQEEDAKKTFKESREAALNLAY  235


RESULT 4
US-09-902-775A-137
; Sequence 137, Application US/09902775A
; Patent No. 6686451
; GENERAL INFORMATION:
;  APPLICANT: Genentech, Inc.
;  APPLICANT:  Ashkenazi, Avi
;  APPLICANT:  Botstein, David
;  APPLICANT:  Desnoyers, Luc
;  APPLICANT:  Eaton, Dan L.
;  APPLICANT:  Ferrara, Napoleone
;  APPLICANT:  Filvaroff, Ellen
;  APPLICANT:  Fong, Sherman
;  APPLICANT:  Gao, Wei-Qiang
;  APPLICANT:  Gerber, Hanspeter
;  APPLICANT:  Gerritsen, Mary E.
;  APPLICANT:  Goddard, A.
;  APPLICANT:  Godowski, Paul J.
;  APPLICANT:  Grimaldi, Christopher J.
;  APPLICANT:  Gurney, Austin L.
;  APPLICANT:  Hillan, Kenneth, J.
;  APPLICANT:  Kljavin, Ivar J.
;  APPLICANT:  Mather, Jennie P.
;  APPLICANT:  Pan, James
;  APPLICANT:  Paoni, Nicholas F.
;  APPLICANT:  Roy, Margaret Ann
;  APPLICANT:  Stewart, Timothy A.
;  APPLICANT:  Tumas, Daniel
;  APPLICANT:  Williams, P. Mickey
;  APPLICANT:  Wood, William, I.
;  TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
;  TITLE OF INVENTION:  Acids Encoding the Same
;  FILE REFERENCE: 10466-14
;  CURRENT APPLICATION NUMBER: US/09/902,775A
;  CURRENT FILING DATE:  2001-07-10
;  PRIOR APPLICATION NUMBER: PCT/US00/04414
;  PRIOR FILING DATE: 2000-02-22
;  PRIOR APPLICATION NUMBER: US 60/143,048
;  PRIOR FILING DATE: 1999-07-07
;  PRIOR APPLICATION NUMBER: US 60/145,698
;  PRIOR FILING DATE: 1999-07-26
;  PRIOR APPLICATION NUMBER: US 60/146,222
;  PRIOR FILING DATE: 1999-07-28
;  PRIOR APPLICATION NUMBER: PCT/US99/20594
;  PRIOR FILING DATE: 1999-09-08
;  PRIOR APPLICATION NUMBER: PCT/US99/20944
;  PRIOR FILING DATE: 1999-09-13
;  PRIOR APPLICATION NUMBER: PCT/US99/21090
;  PRIOR FILING DATE: 1999-09-15
;  PRIOR APPLICATION NUMBER: PCT/US99/21547
;  PRIOR FILING DATE: 1999-09-15
;  PRIOR APPLICATION NUMBER: PCT/US99/23089
;  PRIOR FILING DATE: 1999-10-05
;  PRIOR APPLICATION NUMBER: PCT/US99/28214
;  PRIOR FILING DATE: 1999-11-29
;  PRIOR APPLICATION NUMBER: PCT/US99/28313
;  PRIOR FILING DATE: 1999-11-30
;  PRIOR APPLICATION NUMBER: PCT/US99/28564
;  PRIOR FILING DATE: 1999-12-02
;  PRIOR APPLICATION NUMBER: PCT/US99/28565
;  PRIOR FILING DATE: 1999-12-02
;  PRIOR APPLICATION NUMBER: PCT/US99/30095
;  PRIOR FILING DATE: 1999-12-16
;  PRIOR APPLICATION NUMBER: PCT/US99/30911
;  PRIOR FILING DATE: 1999-12-20
;  PRIOR APPLICATION NUMBER: PCT/US99/30999
;  PRIOR FILING DATE: 1999-12-20
;  PRIOR APPLICATION NUMBER: PCT/US00/00219
;  PRIOR FILING DATE: 2000-01-05
;  NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 137
;   LENGTH: 382
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-902-775A-137

  Query Match             98.7%;  Score 1101;  DB 4;  Length 382;
  Best Local Similarity   96.3%;  Pred. No. 5.2e-107;
  Matches  206;  Conservative    0;  Mismatches    0;  Indels    8;  Gaps    1;

Qy          1 ATGRLLS--------GQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI  52
              |||||||        ||||||||||||||||||||||||||||||||||||||||||||
Db         22 ATGRLLSASDLDLRGGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI  81
```

```
Qy          53 ESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVD 112
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          82 ESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVD 141

Qy         113 EPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGE 172
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         142 EPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGE 201

Qy         173 ETELTTPVLPEETQEEDAKKTFKESREAALNLAY 206
               ||||||||||||||||||||||||||||||||||
Db         202 ETELTTPVLPEETQEEDAKKTFKESREAALNLAY 235


RESULT 5
US-09-638-203-3
; Sequence 3, Application US/09638203
; Patent No. 6602501
; GENERAL INFORMATION:
;  APPLICANT:  Daniel E.H. Afar
;  APPLICANT:  Rene S. Hubert
;  APPLICANT:  Aya Jakobovits
;  APPLICANT:  Arthur B. Raitano
;  TITLE OF INVENTION:  NOVEL C-TYPE LECTIN TRANSMEMBRANE
;  TITLE OF INVENTION:  ANTIGEN EXPRESSED IN HUMAN PROSTATE CANCER AND USES THEREOF
;  FILE REFERENCE:  129.20USU1
;  CURRENT APPLICATION NUMBER: US/09/638,203
;  CURRENT FILING DATE:  2000-08-11
;  PRIOR APPLICATION NUMBER:  60/148,935
;  PRIOR FILING DATE:  1999-08-12
;  NUMBER OF SEQ ID NOS:  47
;  SOFTWARE:  FastSEQ for Windows Version 4.0
; SEQ ID NO 3
;   LENGTH: 260
;   TYPE: PRT
;   ORGANISM: Hamster
US-09-638-203-3

  Query Match             84.3%;  Score 939.5;  DB 4;  Length 260;
  Best Local Similarity   85.8%;  Pred. No. 2.5e-90;
  Matches  175;  Conservative    9;  Mismatches   19;  Indels    1;  Gaps    1;

Qy           4 RLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEK 63
               |||||| ||||||:||||||||||||  :|||||||||||||||||||||:||||:||||
Db           1 RLLSGQLVCRGGTRRPCYKVIYFHDAFQRLNFEEAKEACRRDGGQLVSIETEDEQRLIEK 60

Qy          64 FIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEPSCGSEVCVV 123
               |||||| |||||||||| | || |:||||||||||||| ||||||||||||||||||||
Db          61 FIENLLASDGDFWIGLRRLEVKQVNNTACQDLYAWTDGSTSQFRNWYVDEPSCGSEVCVV 120

Qy         124 MYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPA-VPSREAEGEETELTTPVLP 182
               |||||||| |||| ||||||||||||||||||:||||:  ||     || ||  |||||
Db         121 MYHQPSAPPGIGGSYMFQWNDDRCNMKNNFICKYADEKPSTTPSIRPGGEATEPPTPVLP 180

Qy         183 EETQEEDAKKTFKESREAALNLAY 206
               ||||:|| |:|||||||||||||
Db         181 EETQKEDTKETFKESREAALNLAY 204


RESULT 6
US-09-638-203-2
; Sequence 2, Application US/09638203
; Patent No. 6602501
; GENERAL INFORMATION:
;  APPLICANT:  Daniel E.H. Afar
;  APPLICANT:  Rene S. Hubert
;  APPLICANT:  Aya Jakobovits
;  APPLICANT:  Arthur B. Raitano
;  TITLE OF INVENTION:  NOVEL C-TYPE LECTIN TRANSMEMBRANE
;  TITLE OF INVENTION:  ANTIGEN EXPRESSED IN HUMAN PROSTATE CANCER AND USES THEREOF
;  FILE REFERENCE:  129.20USU1
;  CURRENT APPLICATION NUMBER: US/09/638,203
;  CURRENT FILING DATE:  2000-08-11
;  PRIOR APPLICATION NUMBER:  60/148,935
;  PRIOR FILING DATE:  1999-08-12
;  NUMBER OF SEQ ID NOS:  47
;  SOFTWARE:  FastSEQ for Windows Version 4.0
; SEQ ID NO 2
;   LENGTH: 273
;   TYPE: PRT
;   ORGANISM: Homo Sapiens
US-09-638-203-2

  Query Match             50.4%;  Score 561.5;  DB 4;  Length 273;
  Best Local Similarity   60.1%;  Pred. No. 1e-50;
  Matches  101;  Conservative   25;  Mismatches   35;  Indels    7;  Gaps    3;

Qy           4 RLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEK 63
               |::||| ||    : ||||: ||| | |::|:||: ||  :|| |:|:|:| ||||||
Db          23 RVVSGQKVCFADFKHPCYKMAYFHELSSRVSFQEARLACESEGGVLLSLENEAEQKLIES 82

Qy          64 FIENLLP-----SDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEPSCGS 118
                 ::||        |||||||||| |  : |: | || ||| |:||| ||:|||| |||||||
Db          83 MLQNLTKPGTGISDGDFWIGLWRNGDGQT-SGACPDLYQWSDGSNSQYRNWYTDEPSCGS 141

Qy         119 EVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDE-KPAVP 165
               | ||||||||:|  |:||||::||||||||||:|:||||  |  |  |
Db         142 EKCVVMYHQPTANPGLGGPYLYQWNDDRCNMKHNYICKYEPEINPTAP 189


RESULT 7
US-09-489-847-325
; Sequence 325, Application US/09489847
; Patent No. 6476195
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al
;  TITLE OF INVENTION: 98 Human Secreted Proteins
;  FILE REFERENCE: PZ031P1
;  CURRENT APPLICATION NUMBER: US/09/489,847
;  CURRENT FILING DATE: 2000-01-24
;  EARLIER APPLICATION NUMBER: PCT/US99/17130
;  EARLIER FILING DATE: 1999-07-29
;  EARLIER APPLICATION NUMBER: 60/094,657
;  EARLIER FILING DATE: 1998-07-30
;  EARLIER APPLICATION NUMBER: 60/095,486
;  EARLIER FILING DATE: 1998-08-05
;  EARLIER APPLICATION NUMBER: 60/096,319
;  EARLIER FILING DATE: 1998-08-12
;  EARLIER APPLICATION NUMBER: 60/095,454
;  EARLIER FILING DATE: 1998-08-06
;  EARLIER APPLICATION NUMBER: 60/095,455
;  EARLIER FILING DATE: 1998-08-06
;  NUMBER OF SEQ ID NOS: 376
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 325
;   LENGTH: 81
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-489-847-325

  Query Match             21.5%;  Score 240;  DB 4;  Length 81;
  Best Local Similarity   100.0%;  Pred. No. 9.1e-18;
  Matches   44;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy           1 ATGRLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRR 44
               |||||||||||||||||||||||||||||||||||||||||||
Db          22 ATGRLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRR 65


RESULT 8
US-09-489-847-234
; Sequence 234, Application US/09489847
; Patent No. 6476195
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al
;  TITLE OF INVENTION: 98 Human Secreted Proteins
;  FILE REFERENCE: PZ031P1
```

```
;  CURRENT APPLICATION NUMBER: US/09/489,847
;  CURRENT FILING DATE: 2000-01-24
;  EARLIER APPLICATION NUMBER: PCT/US99/17130
;  EARLIER FILING DATE: 1999-07-29
;  EARLIER APPLICATION NUMBER: 60/094,657
;  EARLIER FILING DATE: 1998-07-30
;  EARLIER APPLICATION NUMBER: 60/095,486
;  EARLIER FILING DATE: 1998-08-05
;  EARLIER APPLICATION NUMBER: 60/096,319
;  EARLIER FILING DATE: 1998-08-12
;  EARLIER APPLICATION NUMBER: 60/095,454
;  EARLIER FILING DATE: 1998-08-06
;  EARLIER APPLICATION NUMBER: 60/095,455
;  EARLIER FILING DATE: 1998-08-06
;  NUMBER OF SEQ ID NOS: 376
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 234
;   LENGTH: 82
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: SITE
;   LOCATION: (82)
;   OTHER INFORMATION: Xaa equals stop translation
US-09-489-847-234

  Query Match             21.5%;  Score 240;  DB 4;  Length 82;
  Best Local Similarity   100.0%;  Pred. No. 9.3e-18;
  Matches   44;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ATGRLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRR 44
              |||||||||||||||||||||||||||||||||||||||||||
Db         22 ATGRLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRR 65


RESULT 9
US-09-489-847-324
; Sequence 324, Application US/09489847
; Patent No. 6476195
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al
;  TITLE OF INVENTION: 98 Human Secreted Proteins
;  FILE REFERENCE: PZ031P1
;  CURRENT APPLICATION NUMBER: US/09/489,847
;  CURRENT FILING DATE: 2000-01-24
;  EARLIER APPLICATION NUMBER: PCT/US99/17130
;  EARLIER FILING DATE: 1999-07-29
;  EARLIER APPLICATION NUMBER: 60/094,657
;  EARLIER FILING DATE: 1998-07-30
;  EARLIER APPLICATION NUMBER: 60/095,486
;  EARLIER FILING DATE: 1998-08-05
;  EARLIER APPLICATION NUMBER: 60/096,319
;  EARLIER FILING DATE: 1998-08-12
;  EARLIER APPLICATION NUMBER: 60/095,454
;  EARLIER FILING DATE: 1998-08-06
;  EARLIER APPLICATION NUMBER: 60/095,455
;  EARLIER FILING DATE: 1998-08-06
;  NUMBER OF SEQ ID NOS: 376
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 324
;   LENGTH: 115
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-489-847-324

  Query Match             21.5%;  Score 240;  DB 4;  Length 115;
  Best Local Similarity   100.0%;  Pred. No. 1.5e-17;
  Matches   44;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ATGRLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRR 44
              |||||||||||||||||||||||||||||||||||||||||||
Db         56 ATGRLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRR 99


RESULT 10
US-09-976-594-168
; Sequence 168, Application US/09976594
; Patent No. 6673549
; GENERAL INFORMATION:
;  APPLICANT: Furness, Michael
;  APPLICANT:  Buchbinder, Jenny
;  TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROIDS
;  FILE REFERENCE: PA-0041 US
;  CURRENT APPLICATION NUMBER: US/09/976,594
;  CURRENT FILING DATE:  2001-10-12
;  PRIOR APPLICATION NUMBER: 60/240,409
;  PRIOR FILING DATE: 2000-10-12
;  NUMBER OF SEQ ID NOS: 1143
;  SOFTWARE: PERL Program
; SEQ ID NO 168
;   LENGTH: 1456
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No. 6673549 1674368CD1
US-09-976-594-168

  Query Match             16.3%;  Score 182;  DB 4;  Length 1456;
  Best Local Similarity   25.5%;  Pred. No. 5.5e-10;
  Matches   52;  Conservative   37;  Mismatches   61;  Indels   54;  Gaps    8;

Qy         21 YKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFWIGLR 80
              ||   :: :  :   : |:  |:|: | ||||:|| |:| : |:: |   :   ::|||
Db        807 YKDYQYYFSKEKETMDNARAFCKRNFGDLVSIQSESEKKFLWKYV-NRNDAQSAYFIGLL 865

Qy         81 RREEKQSNSTACQDLYAWTDGSISQFRNWYVDEPSCGS--EVCVVMYHQPSAPAGIGGPY 138
                 :|:          :|| |||   : :|   ||:  :  | || ||    : :|
Db        866 ISLDKK---------FAWMDGSKVDYVSWATGEPNFANEDENCVTMY----SNSGF---- 908

Qy        139 MFQWNDDRCNMKNNFICKYSDEK----PAVPSREAEGEETELTTPVLPEETQE------- 187
              |||  |   | |||:  :          :|           | | :|   :|
Db        909 ---WNDINCGYPNAFICQRHNSSINATTVMP----------TMPSVPSGCKEGWNFYSN 954

Qy        188 ---------EDAKKTFKESREAAL 202
                          |: :| ::|:|:| :
Db        955 KCFKIFGFMEEERKNWQEARKACI 978


RESULT 11
US-08-840-062-5
; Sequence 5, Application US/08840062
; Patent No. 6117977
;  GENERAL INFORMATION:
;    APPLICANT:  LASKY, LAURENCE A.
;    APPLICANT:  WU, KAI
;    TITLE OF INVENTION:  TYPE C LECTINS
;    NUMBER OF SEQUENCES:  15
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Genentech, Inc.
;      STREET:  460 Point San Bruno Blvd
;      CITY:  South San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94080
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5 inch, 1.44 Mb floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  WinPatin (Genentech)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/840,062
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Dreger, Ginger R.
```

```
;       REGISTRATION NUMBER:  33,055
;       REFERENCE/DOCKET NUMBER:  P1019R1
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  415/225-3216
;       TELEFAX:  415/952-9881
;       TELEX:  910/371-7168
;   INFORMATION FOR SEQ ID NO:  5:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  1455 amino acids
;       TYPE:  Amino Acid
;       TOPOLOGY:  Linear
US-08-840-062-5

  Query Match             16.0%;  Score 178.5;  DB 3;  Length 1455;
  Best Local Similarity   25.9%;  Pred. No. 1.3e-09;
  Matches   53;  Conservative   33;  Mismatches   66;  Indels   53;  Gaps    8;

Qy         21 YKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFWIGLR 80
              ||   :: :  :   : |:  |::: | | :|:|| |:| :.|:| |       ::||:
Db        806 YKDYQYYFSKEKETMDNARRFCKKNFGDLATIKSESEKKFLWKYI-NKNGGQSPYFIGML 864

Qy         81 RREEKQSNSTACQDLYAWTDGSISQFRNWYVDEPSCGS--EVCVVMYHQPSAPAGIGGPY 138
                :|:         : | |||   |  |   ||:  :  | || ||      :|
Db        865 ISMDKK---------FIWMDGSKVDFVAWATGEPNFANDDENCVTMY----TNSGF---- 907

Qy        139 MFQWNDDRCNMKNNFICK---YSDEKPAVPSREAEGEETELTTPVLPEETQE-------- 187
                |||  |   |||||:    |    |:|          |||  |   :|
Db        908 ---WNDINCGYPNNFICQRHNSSINATAMP----------TTPTTPGGCKEGWHLYKNK 953

Qy        188 --------EDAKKTFKESREAALNL 204
                     : ||:::::|:|   |
Db        954 CFKIFGFANEEKKSWQDARQACKGL 978


RESULT 12
US-08-840-062-4
; Sequence 4, Application US/08840062
; Patent No. 6117977
;  GENERAL INFORMATION:
;    APPLICANT:  LASKY, LAURENCE A.
;    APPLICANT:  WU, KAI
;    TITLE OF INVENTION:  TYPE C LECTINS
;    NUMBER OF SEQUENCES:  15
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Genentech, Inc.
;      STREET:  460 Point San Bruno Blvd
;      CITY:  South San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94080
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5 inch, 1.44 Mb floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  WinPatin (Genentech)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/840,062
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Dreger, Ginger R.
;      REGISTRATION NUMBER:  33,055
;      REFERENCE/DOCKET NUMBER:  P1019R1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415/225-3216
;      TELEFAX:  415/952-9881
;      TELEX:  910/371-7168
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1479 amino acids
;      TYPE:  Amino Acid
;      TOPOLOGY:  Linear
US-08-840-062-4

  Query Match             15.8%;  Score 176.5;  DB 3;  Length 1479;
  Best Local Similarity   30.4%;  Pred. No. 2.1e-09;
  Matches   51;  Conservative   25;  Mismatches   57;  Indels   35;  Gaps    7;

Qy          7 SGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIE 66
              | ||       |  ||::      : : :::|:|:|| | || |||| |  | : | | |:
Db        385 SWQPF-----QGHCYRL-----QAEKRSWQESKKACLRGGGDLVSIHSMAELEFITKQIK 434

Qy         67 NLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEPS---CGSEVCVV 123
                :    : ||||        |    |  : |:|||:  | :|:  ||:       | ||
Db        435 QEVE---ELWIGL--------NDLKLQMNFEWSDGSLVSFTHWHPFEPNNFRDSLEDCVT 483

Qy        124 MYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEG 171
                ::     | |       :|||  ||     ||| : :     : |  |
Db        484 IW----GPEG-------RWNDSPCNQSLPSICKKAGQLSQGAAEEDHG 520


RESULT 13
US-08-340-428B-49
; Sequence 49, Application US/08340428B
; Patent No. 5648465
;  GENERAL INFORMATION:
;    APPLICANT:  MARGOLIS, Richard U.
;    APPLICANT:  RAUCH, Uwe
;    APPLICANT:  MARGOLIS, Renee K.
;    TITLE OF INVENTION:  CLONING, EXPRESSION AND USES FOR A
;    TITLE OF INVENTION:  NEUROCAN AS A CHONDROITIN SULFATE PROTEOGLYCAN
;    NUMBER OF SEQUENCES:  49
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Browdy and Neimark
;      STREET:  419 Seventh Street, N.W.
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  U.S.A.
;      ZIP:  20004
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/340,428B
;      FILING DATE:  14 No. 5648465ember 1994
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/922,911
;      FILING DATE:  03 August 1992
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Browdy, Roger L.
;      REGISTRATION NUMBER:  25,618
;      REFERENCE/DOCKET NUMBER:  Margolis=1A
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202-628-5197
;      TELEFAX:  202-737-3528
;  INFORMATION FOR SEQ ID NO:  49:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1257 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-340-428B-49

  Query Match             15.6%;  Score 174;  DB 1;  Length 1257;
  Best Local Similarity   30.8%;  Pred. No. 3.1e-09;
  Matches   44;  Conservative   18;  Mismatches   49;  Indels   32;  Gaps    6;

Qy         17 QRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFW 76
                |  ||:  ||   : |  :|:|:  |||  | | |: | :| | | |       : |
Db       1037 QGHCYR--YF---AHRRAWEDAERDCRRRAGHLTSVHSPEEHKFINSF------GHENSW 1085
```

```
Qy          77 IGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEPS---CGSEVCVVMYHQPSAPAG 133
               |||  |  ::        : ||| :  |: ||  :|     | | ||||    :
Db        1086 IGLNDRTVERD--------FQWTDNTGLQYENWREKQPDNFFAGGEDCVVMVAHENG--- 1134

Qy          134 IGGPYMFQWNDDRCNMKNNFICK 156
                 :|||  ||    ::||
Db        1135 -------RWNDVPCNYNLPYVCK 1150


RESULT 14
5180808-2
;Patent No. 5180808
;    APPLICANT: RUOSLAHTI, ERKKI I.
;    TITLE OF INVENTION: VERSICAN CORE PROTEIN, NUCLEIC ACID
;SEQUENCES ENCODING THE SAME, NUCLEIC ACID PROBES, ANTI-VERSICAN
;ANTIBODIES, AND METHODS OF DETECTING THE SAME
;    NUMBER OF SEQUENCES: 4
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/441,179
;      FILING DATE: 27-NOV-1989
;SEQ ID NO:2:
;      LENGTH: 2409
5180808-2

  Query Match             15.6%;  Score 174;  DB 6;  Length 2409;
  Best Local Similarity   28.5%;  Pred. No. 7.5e-09;
  Matches   47;  Conservative   23;  Mismatches   55;  Indels   40;  Gaps    8;

Qy          17 QRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDF- 75
               |  |||  ||   : |  :: |:  ||  |  | || | :||  : :             |:
Db        2190 QGQCYK--YF---AHRRTWDAAERECRLQGAHLTSILSHEEQMFVNRV-------GHDYQ 2237

Qy          76 WIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEP----SCGSEVCVVMYHQPSAP 131
               ||||        |    :  : |||||  |: ||  ::|    | | :  |:::|:
Db        2238 WIGL--------NDKMFEHDFRWTDGSTLQYENWRPNQPDSFFSAGEDCVVIIWHENG-- 2287

Qy         132 AGIGGPYMFQWNDDRCNMKNNFICKYS----DEKPAVPSREAEGE 172
                  ||||  ||    : ||       : | | : :  |:
Db        2288 ---------QWNDVPCNYHLTYTCKKGTVACGQPPVVENAKTFGK 2323


RESULT 15
US-08-840-062-2
; Sequence 2, Application US/08840062
; Patent No. 6117977
;  GENERAL INFORMATION:
;    APPLICANT:  LASKY, LAURENCE A.
;    APPLICANT:  WU, KAI
;    TITLE OF INVENTION:  TYPE C LECTINS
;    NUMBER OF SEQUENCES:  15
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Genentech, Inc.
;      STREET:  460 Point San Bruno Blvd
;      CITY:  South San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94080
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5 inch, 1.44 Mb floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  WinPatin (Genentech)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/840,062
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Dreger, Ginger R.
;      REGISTRATION NUMBER:  33,055
;      REFERENCE/DOCKET NUMBER:  P1019R1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415/225-3216
;      TELEFAX:  415/952-9881
;      TELEX:  910/371-7168
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1479 amino acids
;      TYPE:  Amino Acid
;      TOPOLOGY:  Linear
US-08-840-062-2

  Query Match             15.3%;  Score 170.5;  DB 3;  Length 1479;
  Best Local Similarity   31.4%;  Pred. No. 9e-09;
  Matches   48;  Conservative   22;  Mismatches   48;  Indels   35;  Gaps    7;

Qy           7 SGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIE 66
               | ||      |  ||::      : : :::|:| || | || |:|| |  | : | | |:
Db         384 SWQPF-----QGHCYRL-----QAEKRSWQESKRACLRGGGDLLSIHSMAELEFITKQIK 433

Qy          67 NLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEPS---CGSEVCVV 123
                 :    : ||||        |    |  : |:|||:  | :|:  ||:      | ||
Db         434 QEVE---ELWIGL--------NDLKLQMNFEWSDGSLVSFTHWHPFEPNNFRDSLEDCVT 482

Qy         124 MYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICK 156
               ::     | |       :|||  ||      |||
Db         483 IW----GPEG-------RWNDSPCNQSLPSICK 504


RESULT 16
PCT-US95-03747-2
; Sequence 2, Application PC/TUS9503747
;  GENERAL INFORMATION:
;    APPLICANT:  LA JOLLA CANCER RESEARCH FOUNDATION
;    TITLE OF INVENTION:  Brevican, A Glial Cell Proteoglycan
;    NUMBER OF SEQUENCES:  3
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Campbell and Flores
;      STREET:  4370 La Jolla Village Drive, Suite 700
;      CITY:  San Diego
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  92122
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US95/03747
;      FILING DATE:  27-MAR-1995
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Imbra, Richard J.
;      REGISTRATION NUMBER:  37,643
;      REFERENCE/DOCKET NUMBER:  FP-LJ 1453
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (619) 535-9001
;      TELEFAX:  (619) 535-8949
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  912 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
PCT-US95-03747-2

  Query Match             13.6%;  Score 152;  DB 5;  Length 912;
  Best Local Similarity   28.8%;  Pred. No. 4e-07;
  Matches   47;  Conservative   17;  Mismatches   61;  Indels   38;  Gaps    8;

Qy          17 QRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFW 76
               |  |||    | ::|| ::|||:  ||  |  | || : :||  |                |
Db         697 QGACYK----HFSARR-SWEEAENKCRMYGAHLASISTPEEQDFINNRYREYQ------W 745

Qy          77 IGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEPS---CGSEVCVVM-YHQPSAPA 132
               |||        |    :  : |:||    : ||   :|       | ||||  :|
```

```
Db        746 IGL--------NDRTIEGDFLWSDGVPLLYENWNPGQPDSYFLSGENCVVMVWHDQG--- 794

Qy        133 GIGGPYMFQWNDDRCNMKNNFICKYS----DEKPAVPSREAEG 171
                ||:|  ||   :: ||         | :|  |  |
Db        795 --------QWSDVPCNYHLSYTCKMGLVSCGPPPELPLAEVFG 829


RESULT 17
US-09-602-877A-99
; Sequence 99, Application US/09602877A
; Patent No. 6432707
; GENERAL INFORMATION:
;  APPLICANT:  Reed, Steven G.
;  APPLICANT:  Xu, Jiangchun
;  APPLICANT:  Dillon, Davin C.
;  TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF BREAST CANCER
;  FILE REFERENCE:  210121.446C5
;  CURRENT APPLICATION NUMBER: US/09/602,877A
;  CURRENT FILING DATE:  2000-06-22
;  NUMBER OF SEQ ID NOS:  107
;  SOFTWARE:  FastSEQ for Windows Version 3.0
; SEQ ID NO 99
;   LENGTH: 197
;   TYPE: PRT
;   ORGANISM: Homo sapien
US-09-602-877A-99

  Query Match             13.5%;  Score 151;  DB 4;  Length 197;
  Best Local Similarity   25.0%;  Pred. No. 6.4e-08;
  Matches   38;  Conservative   26;  Mismatches   64;  Indels   24;  Gaps    4;

Qy          9 QPVCRGGTQ--RPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIE 66
              | ||  ||:  : ||       :    :| || | |   || ||   : ||   :: : :
Db         65 QTVCLRGTKVHKKCYLA-----SEGLKHFHEANEDCISKGGILVIPRNSDEINALQDYGK 119

Qy         67 NLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEPSCGSEVCVVMYH 126
                ||   |||:|:        |    :  :   :|    | ||   :|: |     |::
Db        120 RSLPGVNDFWLGI--------NDMVTEGKFVDVNGIAISFLNWDRAQPNGGKRENCVLFS 171

Qy        127 QPSAPAGIGGPYMFQWNDDRCNMKNNFICKYS 158
              | :          :|:|: |     :||:::
Db        172 QSA---------QGKWSDEACRSSKRYICEFT 194


RESULT 18
US-09-866-028-50
; Sequence 50, Application US/09866028
; Patent No. 6642360
; GENERAL INFORMATION:
;  APPLICANT: Baker, Kevin
;  APPLICANT:  Botstein,David
;  APPLICANT:  Eaton,Dan
;  APPLICANT:  Ferrara,Napoleone
;  APPLICANT:  Filvaroff,Ellen
;  APPLICANT:  Gerritsen,Mary
;  APPLICANT:  Goddard,Audrey
;  APPLICANT:  Godowski,Paul
;  APPLICANT:  Grimaldi,Christopher
;  APPLICANT:  Gurney,Austin
;  APPLICANT:  Hillan,Kenneth
;  APPLICANT:  Kljavin,Ivar
;  APPLICANT:  Napier,Mary
;  APPLICANT:  Roy,Margaret
;  APPLICANT:  Tumas,Daniel
;  APPLICANT:  Wood,William
;  TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
;  TITLE OF INVENTION:  ACIDS ENCODING THE SAME
;  FILE REFERENCE: P2548P1C1
;  CURRENT APPLICATION NUMBER: US/09/866,028
;  CURRENT FILING DATE:  2001-05-25
;  Prior application data removed - consult PALM or file wrapper
;  NUMBER OF SEQ ID NOS: 120
; SEQ ID NO 50
;   LENGTH: 455
;   TYPE: PRT
;   ORGANISM: Homo Sapien
US-09-866-028-50

  Query Match             13.4%;  Score 149.5;  DB 4;  Length 455;
  Best Local Similarity   25.6%;  Pred. No. 2.9e-07;
  Matches   45;  Conservative   25;  Mismatches   61;  Indels   45;  Gaps    8;

Qy         14 GGTQRPCYKVIY--FHDTSRRLN-----------FEEAKEACRRDGGQLVSIESEDEQK 59
              || |  |   ::  ||     |::            :  |:  |:| || |   |:|:  |
Db        287 GGAQ--CATKVHFPFHTCDLRIDGDCFMVSSEADTYYRARMKCQRKGGVLAQIKSQKVQD 344

Qy         60 LIEKFIENLLP----SDGD-----FWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWY 110
              ::  ::  |       :| |     |||||  :  |         | : |  |     | ::
Db        345 ILAFYLGRLETTNEVTDSDFETRNFWIGLTYKTAK--------DSFRWATGEHQAFTSFA 396

Qy        111 VDEPSCGSEVCVVMYHQPSAPAGIG------GPYMFQWNDDRCNMKNNFICKYSDE 160
                :|     | :      ||  | |          | ||| ||  :| :||::: |
Db        397 FGQPDNHGLVWL------SAAMGFGNCVELQASAAFNWNDQRCKTRNRYICQFAQE 446


RESULT 19
US-07-641-971B-1
; Sequence 1, Application US/07641971B
; Patent No. 5236706
;  GENERAL INFORMATION:
;    APPLICANT:  Debre, Patrice
;    APPLICANT:  Mossalayi, Mohammed D
;    TITLE OF INVENTION:  A PHARMACEUTICAL PREPARATION FOR THE
;    TITLE OF INVENTION:  MATURATION OF PROTHYMOCYTES
;    NUMBER OF SEQUENCES:  6
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Irving M. Fishman, CIBA-GEIGY Corporation
;      STREET:  556 Morris Avenue
;      CITY:  Summit
;      STATE:  New Jersey
;      COUNTRY:  USA
;      ZIP:  07901
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/641,971B
;      FILING DATE:  19910116
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  GB 90016254
;      FILING DATE:  24-JAN-1990
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Fishman, Irving M
;      REGISTRATION NUMBER:  30258
;      REFERENCE/DOCKET NUMBER:  4-17921/+/DEB
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  908-277-4832
;      TELEFAX:  908-277-4306
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  174 amino acids
;      TYPE:  AMINO ACID
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    FRAGMENT TYPE:  N-terminal
;    ORIGINAL SOURCE:
;      CELL TYPE:  Human B. Cells
;      CELL LINE:  CHO cells transformed with pCAL8-BF-ND
US-07-641-971B-1
```

```
  Query Match             13.4%;  Score 149;  DB 1;  Length 174;
  Best Local Similarity   30.2%;  Pred. No. 8.7e-08;
  Matches   55;  Conservative   20;  Mismatches   57;  Indels   50;  Gaps   11;

Qy         17 QRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFW 76
              || ||   ||   :::  :  |: ||    |||||| | :|| : |       |    |
Db         24 QRKCY---YFGKGTKQ--WVHARYACDDMEGQLVSIHSPEEQDFLTKH-----ASHTGSW 73

Qy         77 IGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEPSCGS--EVCVVMYHQPSAPAGI 134
              ||||  : |          : | |||   : ||   ||:  |  | ||:|       |
Db         74 IGLRNLDLKGE--------FIWVDGSHVDYSNWAPGEPTSRSQGEDCVMM-------RGS 118

Qy        135 GGPYMFQWNDDRCNMK-NNFICKYSDEKPAV---PSREAEGE----------ETELTTPV 180
              |     :|||  |: |   ::|    :: |    |: |   |          : | ||
Db        119 G-----RWNDAFCDRKLGAWVC----DRLATCTPPASEGSAESMGPDSRPDPDGRLPTPS 169

Qy        181 LP 182
               |
Db        170 AP 171
```

RESULT 20
US-07-781-248A-1
; Sequence 1, Application US/07781248A
; Patent No. 5246699
; GENERAL INFORMATION:
; APPLICANT: Debre, Patrice
; APPLICANT: Mossalayi, Mohammed D
; TITLE OF INVENTION: MATURATION OF HEMATOPOIETIC CELLS
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Irving M. Fishman, CIBA-GEIGY Corporation
; STREET: 556 Morris Avenue
; CITY: Summit
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07901
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/781,248A
; FILING DATE: 19911230
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 90103565
; FILING DATE: 09-MAY-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Ikeler, Barbara J.
; REGISTRATION NUMBER: 36,170
; REFERENCE/DOCKET NUMBER: 4-18065/A/DEB
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 908-277-3368
; TELEFAX: 908-277-4306
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 174 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; CELL TYPE: Human B. Cells
; CELL LINE: CHO cells transformed with pCAL8-BF-ND
US-07-781-248A-1


```
  Query Match             13.4%;  Score 149;  DB 1;  Length 174;
  Best Local Similarity   30.2%;  Pred. No. 8.7e-08;
  Matches   55;  Conservative   20;  Mismatches   57;  Indels   50;  Gaps   11;
```


```
Qy         17 QRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFW 76
              || ||   ||   :::  :  |: ||    |||||| | :|| : |       |    |
Db         24 QRKCY---YFGKGTKQ--WVHARYACDDMEGQLVSIHSPEEQDFLTKH-----ASHTGSW 73

Qy         77 IGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEPSCGS--EVCVVMYHQPSAPAGI 134
              ||||  : |          : | |||   : ||   ||:  |  | ||:|       |
Db         74 IGLRNLDLKGE--------FIWVDGSHVDYSNWAPGEPTSRSQGEDCVMM-------RGS 118

Qy        135 GGPYMFQWNDDRCNMK-NNFICKYSDEKPAV---PSREAEGE----------ETELTTPV 180
              |     :|||  |: |   ::|    :: |    |: |   |          : | ||
Db        119 G-----RWNDAFCDRKLGAWVC----DRLATCTPPASEGSAESMGPDSRPDPDGRLPTPS 169

Qy        181 LP 182
               |
Db        170 AP 171
```

RESULT 21
US-08-365-103B-10
; Sequence 10, Application US/08365103B
; Patent No. 5766943
; GENERAL INFORMATION:
; APPLICANT: Lynch, Richard G
; APPLICANT: Nunez, Raphael D.
; APPLICANT: Yodoi, Jungi
; TITLE OF INVENTION: DNA Sequences for Soluble Froms of CD23
; TITLE OF INVENTION: and Methods of Use for Same
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Zarley, McKee, Thomte, Voorhees & Sease
; STREET: 801 Grand Ave. Suite 3200
; CITY: Des Moines
; STATE: Iowa
; COUNTRY: United States
; ZIP: 50309
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/365,103B
; FILING DATE: 28-DEC-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Nebel, Heidi S.
; REGISTRATION NUMBER: 37,719
; REFERENCE/DOCKET NUMBER: Uirf N5-24
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (515) 288-3667
; TELEFAX: (515) 288-1338
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 320 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-365-103B-10

```
  Query Match             13.4%;  Score 149;  DB 1;  Length 320;
  Best Local Similarity   30.2%;  Pred. No. 2e-07;
  Matches   55;  Conservative   20;  Mismatches   57;  Indels   50;  Gaps   11;

Qy         17 QRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFW 76
              || ||   ||   :::  :  |: ||    |||||| | :|| : |       |    |
Db        170 QRKCY---YFGKGTKQ--WVHARYACDDMEGQLVSIHSPEEQDFLTKH-----ASHTGSW 219

Qy         77 IGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEPSCGS--EVCVVMYHQPSAPAGI 134
              ||||  : |          : | |||   : ||   ||:  |  | ||:|       |
Db        220 IGLRNLDLKGE--------FIWVDGSHVDYSNWAPGEPTSRSQGEDCVMM-------RGS 264

Qy        135 GGPYMFQWNDDRCNMK-NNFICKYSDEKPAV---PSREAEGE----------ETELTTPV 180
```

```
Db        265 G-----RWNDAFCDRKLGAWVC----DRLATCTPPASEGSAESMGPDSRPDPDGRLPTPS 315

Qy        181 LP 182
               |
Db        316 AP 317


RESULT 22
US-08-365-103B-8
; Sequence 8, Application US/08365103B
; Patent No. 5766943
;  GENERAL INFORMATION:
;    APPLICANT:  Lynch, Richard G
;    APPLICANT:  Nunez, Raphael D.
;    APPLICANT:  Yodoi, Jungi
;    TITLE OF INVENTION:  DNA Sequences for Soluble Froms of CD23
;    TITLE OF INVENTION:  and Methods of Use for Same
;    NUMBER OF SEQUENCES:  14
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Zarley, McKee, Thomte, Voorhees & Sease
;      STREET:  801 Grand Ave. Suite 3200
;      CITY:  Des Moines
;      STATE:  Iowa
;      COUNTRY:  United States
;      ZIP:  50309
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/365,103B
;      FILING DATE:  28-DEC-1994
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Nebel, Heidi S.
;      REGISTRATION NUMBER:  37,719
;      REFERENCE/DOCKET NUMBER:  Uirf N5-24
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (515) 288-3667
;      TELEFAX:  (515) 288-1338
;  INFORMATION FOR SEQ ID NO:  8:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  321 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-365-103B-8

  Query Match             13.4%;  Score 149;  DB 1;  Length 321;
  Best Local Similarity   30.2%;  Pred. No. 2e-07;
  Matches   55;  Conservative   20;  Mismatches   57;  Indels   50;  Gaps   11;

Qy         17 QRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFW 76
              || ||   ||   :::  :  |: ||    |||||| | :||  : |       |    |
Db        171 QRKCY---YFGKGTKQ--WVHARYACDDMEGQLVSIHSPEEQDFLTKH-----ASHTGSW 220

Qy         77 IGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEPSCGS--EVCVVMYHQPSAPAGI 134
              ||||  : |          : | |||   : ||   ||:  |  | ||:|        |
Db        221 IGLRNLDLKGE--------FIWVDGSHVDYSNWAPGEPTSRSQGEDCVMM-------RGS 265

Qy        135 GGPYMFQWNDDRCNMK-NNFICKYSDEKPAV---PSREAEGE----------ETELTTPV 180
              |     :|||  |: |   ::|    :: |    |: |   |          :  | ||
Db        266 G-----RWNDAFCDRKLGAWVC----DRLATCTPPASEGSAESMGPDSRPDPDGRLPTPS 316

Qy        181 LP 182
               |
Db        317 AP 318


RESULT 23
US-08-751-305-2
; Sequence 2, Application US/08751305
; Patent No. 5965439
;  GENERAL INFORMATION:
;    APPLICANT:  Tenner et al., Andrea J.
;    TITLE OF INVENTION:  HOST DEFENSE ENHANCEMENT
;    NUMBER OF SEQUENCES:  33
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fish & Richardson P.C.
;      STREET:  4225 Executive Square, Suite 1400
;      CITY:  La Jolla
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  92037
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/751,305
;      FILING DATE:  18-NOV-1996
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Wetherell, Jr., John R.
;      REGISTRATION NUMBER:  31,678
;      REFERENCE/DOCKET NUMBER:  07306/012001
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  619/678-5070
;      TELEFAX:  619/678-5099
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  652 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-751-305-2

  Query Match             13.3%;  Score 148.5;  DB 2;  Length 652;
  Best Local Similarity   24.4%;  Pred. No. 5.9e-07;
  Matches   50;  Conservative   37;  Mismatches   75;  Indels   43;  Gaps    9;

Qy          5 LLSGQP-----------VCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIE 53
              ||  ||           || |     ||       | :|:  ||:  | ::|| | :::
Db         13 LLLTQPGAGTGADTEAVVCVG---TACYTA-----HSGKLSAAEAQNHCNQNGGNLATVK 64

Qy         54 SEDEQKLIEKFIENLLPSD-------GDFWIGLRRREEKQSNSTACQDLYAWT-DGSISQ 105
              |::| : ::: :  ||  :         |||||:| : |  : :     ::|   |  :
Db         65 SKEEAQHVQRVLAQLLRREAALTARMSKFWIGLQREKGKCLDPSLPLKGFSWVGGGEDTP 124

Qy        106 FRNWYVD-EPSCGSEVCVVM---YHQPSAPAGIGGPYMFQWNDDRCNMKNN-------FI 154
              : ||: :   || |: || :      ||  |  :       :|::  |    :       |:
Db        125 YSNWHKELRNSCISKRCVSLLLDLSQPLLPNRLP-----KWSEGPCGSPGSPGSNIEGFV 179

Qy        155 CKYSDEKPAVPSREAEGEETELTTP 179
              ||:| :    |       :  |||
Db        180 CKFSFKGMCRPLALGGPGQVTYTTP 204


RESULT 24
US-09-517-605-2
; Sequence 2, Application US/09517605
; Patent No. 6391567
; GENERAL INFORMATION:
;  APPLICANT: Littman, Dan R.
;  APPLICANT: Kwon, Douglas S.
;  APPLICANT: van Kooyk, Yvette
;  APPLICANT: Geijtenbeck, Tneo
;  TITLE OF INVENTION: METHODS OF USING A FACILITATOR OF RETROVIRAL ENTRY INTO
;  TITLE OF INVENTION: CELLS
;  FILE REFERENCE: 1049-1-017
;  CURRENT APPLICATION NUMBER: US/09/517,605
;  CURRENT FILING DATE: 2000-03-02
;  NUMBER OF SEQ ID NOS: 17
```

```
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
;   LENGTH: 404
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-517-605-2

  Query Match             13.0%;  Score 144.5;  DB 4;  Length 404;
  Best Local Similarity   29.7%;  Pred. No. 8.1e-07;
  Matches   47;  Conservative   23;  Mismatches   47;  Indels   41;  Gaps   10;

Qy         25 YFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFWIGLRRREE 84
              ||   |:| |: ::  ||:  | ||| |:| :||  ::        |:   |:||    :
Db        268 YFMSNSQR-NWHDSITACKEVGAQLVVIKSAEEQNFLQLQSSR---SNRFTWMGLSDLNQ 323

Qy         85 KQSNSTACQDLYAWTDGS--ISQFRN-WYVDEP-SCGSEVCVVMYHQPSAPAGIGGPYMF 140
                    :  : | |||  :  |:  |   || : | | |          :| |
Db        324 --------EGTWQWVDGSPLLPSFKQYWNRGEPNNVGEEDCAEF-------SGNG----- 363

Qy        141 QWNDDRCNMKNNFICKYS------DEK------PAVPS 166
               ||||:||:  :||| |      ||:      || |:
Db        364 -WNDDKCNLAKFWICKKSAASCSRDEEQFLSPAPATPN 400


RESULT 25
US-08-513-278-4
; Sequence 4, Application US/08513278
; Patent No. 5840844
;  GENERAL INFORMATION:
;    APPLICANT:  LASKY, LAURENCE A.
;    APPLICANT:  STACHELL, SCOTT E.
;    APPLICANT:  ROSEN, STEVEN D.
;    APPLICANT:  SINGER, MARK S.
;    APPLICANT:  YEDNOCK, TED A.
;    TITLE OF INVENTION:  LYMPHOCYTE HOMING RECEPTORS
;    NUMBER OF SEQUENCES:  6
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Genentech, Inc.
;      STREET:  460 Point San Bruno Blvd
;      CITY:  South San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94080
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  5.25 inch, 360 Kb floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  patin (Genentech)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/513,278
;      FILING DATE:  10-AUG-1995
;      CLASSIFICATION:  5530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/059027
;      FILING DATE:  06-MAY-1993
;      APPLICATION NUMBER:  07/786149
;      FILING DATE:  31-OCT-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/315015
;      FILING DATE:  23-FEB-1989
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Dreger, Ginger R.
;      REGISTRATION NUMBER:  33,055
;      REFERENCE/DOCKET NUMBER:  565D1C1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415/225-3216
;      TELEFAX:  415/952-9881
;      TELEX:  910/371-7168
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  372 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear

US-08-513-278-4

  Query Match             12.9%;  Score 143.5;  DB 2;  Length 372;
  Best Local Similarity   26.5%;  Pred. No. 9.2e-07;
  Matches   41;  Conservative   37;  Mismatches   50;  Indels   27;  Gaps    8;

Qy         26 FHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLP-SDGDFWIGLRRREE 84
              :| : : :|:| |:: |:::   ||:|::: |   || ::|| || |   :|||:|:
Db         41 YHYSEKPMNWENARKFCKQNYTDLVAIQNKRE---IE-YLENTLPKSPYYYWIGIRK--- 93

Qy         85 KQSNSTACQDLYAW--TDGSIS-QFRNWYVDEPS--CGSEVCVVMYHQPSAPAGIGGPYM 139
                        :: |  |: ::: :  ||   ||:     | || :| :    :|
Db         94 -------IGKMWTWVGTNKTLTKEAENWGAGEPNNKKSKEDCVEIYIKRERDSG------ 140

Qy        140 FQWNDDRCNMKNNFICKYSDEKPAVPSREAEGEET 174
               :|||| |: :   :|  :  :|   :   |  ||
Db        141 -KWNDDACHKRKAALCYTASCQPGSCNGRGECVET 174


RESULT 26
5514582-4
;Patent No. 5514582
;    APPLICANT: CAPON, DANIEL J.;LASKY, LAURENCE A.
;    TITLE OF INVENTION: RECOMBINANT DNA ENCODING HYBRID
;IMMUNOGLOBULINS
;    NUMBER OF SEQUENCES: 43
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/185,670
;      FILING DATE: 21-JAN-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER: 986,931
;      FILING DATE: 08-DEC-1992
;      APPLICATION NUMBER: 808,122
;      FILING DATE: 16-DEC-1991
;      APPLICATION NUMBER: 440,625
;      FILING DATE: 22-NOV-1989
;      APPLICATION NUMBER: 315,015
;      FILING DATE: 23-FEB-1989
;SEQ ID NO:4:
;      LENGTH: 372
5514582-4

  Query Match             12.9%;  Score 143.5;  DB 6;  Length 372;
  Best Local Similarity   26.5%;  Pred. No. 9.2e-07;
  Matches   41;  Conservative   37;  Mismatches   50;  Indels   27;  Gaps    8;

Qy         26 FHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLP-SDGDFWIGLRRREE 84
              :| : : :|:| |:: |:::   ||:|::: |   || ::|| || |   :|||:|:
Db         41 YHYSEKPMNWENARKFCKQNYTDLVAIQNKRE---IE-YLENTLPKSPYYYWIGIRK--- 93

Qy         85 KQSNSTACQDLYAW--TDGSIS-QFRNWYVDEPS--CGSEVCVVMYHQPSAPAGIGGPYM 139
                        :: |  |: ::: :  ||   ||:     | || :| :    :|
Db         94 -------IGKMWTWVGTNKTLTKEAENWGAGEPNNKKSKEDCVEIYIKRERDSG------ 140

Qy        140 FQWNDDRCNMKNNFICKYSDEKPAVPSREAEGEET 174
               :|||| |: :   :|  :  :|   :   |  ||
Db        141 -KWNDDACHKRKAALCYTASCQPGSCNGRGECVET 174


RESULT 27
US-08-840-062-7
; Sequence 7, Application US/08840062
; Patent No. 6117977
;  GENERAL INFORMATION:
;    APPLICANT:  LASKY, LAURENCE A.
;    APPLICANT:  WU, KAI
;    TITLE OF INVENTION:  TYPE C LECTINS
;    NUMBER OF SEQUENCES:  15
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Genentech, Inc.
;      STREET:  460 Point San Bruno Blvd
;      CITY:  South San Francisco
;      STATE:  California
```

```
;       COUNTRY:  USA
;       ZIP:  94080
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  3.5 inch, 1.44 Mb floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  WinPatin (Genentech)
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/840,062
;       FILING DATE:
;       CLASSIFICATION:  435
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Dreger, Ginger R.
;       REGISTRATION NUMBER:  33,055
;       REFERENCE/DOCKET NUMBER:  P1019R1
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  415/225-3216
;       TELEFAX:  415/952-9881
;       TELEX:  910/371-7168
;   INFORMATION FOR SEQ ID NO:  7:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  1487 amino acids
;       TYPE:  Amino Acid
;       TOPOLOGY:  Linear
US-08-840-062-7

  Query Match             12.7%;  Score 142;  DB 3;  Length 1487;
  Best Local Similarity   26.5%;  Pred. No. 8.7e-06;
  Matches   40;  Conservative   22;  Mismatches   51;  Indels   38;  Gaps    6;

Qy         13 RGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLPSD 72
              | |: | ||:      :    |:: :|  :|   || |:|| |||:  | | :  ::
Db        238 RNGSSRICYQ----FNLLSSLSWNQAHSSCLMQGGALLSIADEDEEDFIRKHLSKVVK-- 291

Qy         73 GDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVD-------EPSCGSEVCVVMY 125
               : |||| : :||          : |:||:   : ||  :       |  ||:   |
Db        292 -EVWIGLNQLDEKAG--------WQWSDGTPLSYLNWSQEITPGPFVEHHCGTLEVV--- 339

Qy        126 HQPSAPAGIGGPYMFQWNDDRCNMKNNFICK 156
                ||           |    |     :|||
Db        340 ---SA----------AWRSRDCESTLPYICK 357


RESULT 28
US-08-722-126A-7
; Sequence 7, Application US/08722126A
; Patent No. 6034227
;  GENERAL INFORMATION:
;    APPLICANT:  PECHT, Israel
;    APPLICANT:  GUTHMANN, Marcelo D.
;    APPLICANT:  TAL, Michael
;    TITLE OF INVENTION:  A DNA MOLECULE ENCODING A MAST CELL
;    TITLE OF INVENTION:  FUNCTION-ASSOCIATED ANTIGEN (MAFA)
;    NUMBER OF SEQUENCES:  20
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  BROWDY AND NEIMARK, P.L.L.C.
;      STREET:  419 Seventh Street N.W., Ste. 300
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  UNITED STATES OF AMERICA
;      ZIP:  20004
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/722,126A
;      FILING DATE:  08-OCT-1996
;      CLASSIFICATION:  536
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US95/04258
;      FILING DATE:  06-APR-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  IL 109257
;      FILING DATE:  08-APR-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  BROWDY, Roger L.
;      REGISTRATION NUMBER:  25,618
;      REFERENCE/DOCKET NUMBER:  PECHT=1A
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 628-5197
;      TELEFAX:  (202) 737-3528
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  125 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-722-126A-7

  Query Match             12.7%;  Score 141.5;  DB 3;  Length 125;
  Best Local Similarity   31.8%;  Pred. No. 3.4e-07;
  Matches   47;  Conservative   21;  Mismatches   37;  Indels   43;  Gaps    9;

Qy         17 QRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFW 76
              |: ||    ||    |::  : :|: ||     |:|||| |: ||  : : |      :  | |
Db          9 QQKCY---YFGKGSKQ--WIQARFACSDLQGRLVSIHSQKEQDFLMQHI-----NKKDSW 58

Qy         77 IGLRRREEKQSNSTACQDL-----YAWTDGSISQFRNWYVDEPSCG--SEVCVVMYHQPS 129
              |||              |||     : |:|||   : ||   ||: |   | ||:|
Db         59 IGL-------------QDLNMEGEFVWSDGSPVGYSNWNPGEPNNGGQGEDCVMM----- 100

Qy        130 APAGIGGPYMFQWNDDRC-NMKNNFICK 156
                 | |     ||||  | :  : ::|:
Db        101 --RGSG-----QWNDAFCRSYLDAWVCE 121


RESULT 29
PCT-US95-04258-7
; Sequence 7, Application PC/TUS9504258
;  GENERAL INFORMATION:
;    APPLICANT:
;    TITLE OF INVENTION:  A DNA MOLECULE ENCODING A MAST CELL
;    TITLE OF INVENTION:  FUNCTION-ASSOCIATED ANTIGEN (MAFA)
;    NUMBER OF SEQUENCES:  10
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  BROWDY AND NEIMARK
;      STREET:  419 Seventh Street, N.W., Suite 300
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20004
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30 (EPO)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US95/04258
;      FILING DATE:  06-APR-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  IL 109257
;      FILING DATE:  08-APR-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  BROWDY, Roger L.
;      REGISTRATION NUMBER:  25,618
;      REFERENCE/DOCKET NUMBER:  PECHT=1 PCT
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202-628-5197
;      TELEFAX:  202-737-3528
;      TELEX:  248633
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  125 amino acids
```

```
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
PCT-US95-04258-7

  Query Match            12.7%;  Score 141.5;  DB 5;  Length 125;
  Best Local Similarity  31.8%;  Pred. No. 3.4e-07;
  Matches   47;  Conservative   21;  Mismatches   37;  Indels   43;  Gaps    9;

Qy         17 QRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFW 76
              |: ||   ||   |::  : :|: ||    |:|||| |: ||  : : |     :  | |
Db          9 QQKCY---YFGKGSKQ--WIQARFACSDLQGRLVSIHSQKEQDFLMQHI-----NKKDSW 58

Qy         77 IGLRRREEKQSNSTACQDL-----YAWTDGSISQFRNWYVDEPSCG--SEVCVVMYHQPS 129
              |||              |||     : |:|||   : ||   ||: |   | ||:|
Db         59 IGL-------------QDLNMEGEFVWSDGSPVGYSNWNPGEPNNGGQGEDCVMM----- 100

Qy        130 APAGIGGPYMFQWNDDRC-NMKNNFICK 156
                | |     ||||  | :  : ::|:
Db        101 --RGSG-----QWNDAFCRSYLDAWVCE 121


RESULT 30
US-08-365-103B-4
; Sequence 4, Application US/08365103B
; Patent No. 5766943
;  GENERAL INFORMATION:
;    APPLICANT:  Lynch, Richard G
;    APPLICANT:  Nunez, Raphael D.
;    APPLICANT:  Yodoi, Jungi
;    TITLE OF INVENTION:  DNA Sequences for Soluble Froms of CD23
;    TITLE OF INVENTION:  and Methods of Use for Same
;    NUMBER OF SEQUENCES:  14
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Zarley, McKee, Thomte, Voorhees & Sease
;      STREET:  801 Grand Ave. Suite 3200
;      CITY:  Des Moines
;      STATE:  Iowa
;      COUNTRY:  United States
;      ZIP:  50309
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/365,103B
;      FILING DATE:  28-DEC-1994
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Nebel, Heidi S.
;      REGISTRATION NUMBER:  37,719
;      REFERENCE/DOCKET NUMBER:  Uirf N5-24
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (515) 288-3667
;      TELEFAX:  (515) 288-1338
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  287 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-365-103B-4

  Query Match            12.7%;  Score 141.5;  DB 1;  Length 287;
  Best Local Similarity  31.8%;  Pred. No. 1e-06;
  Matches   47;  Conservative   21;  Mismatches   37;  Indels   43;  Gaps    9;

Qy         17 QRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFW 76
              |: ||   ||   |::  : :|: ||    |:|||| |: ||  : : |     :  | |
Db        154 QQKCY---YFGKGSKQ--WIQARFACSDLQGRLVSIHSQKEQDFLMQHI-----NKKDSW 203

Qy         77 IGLRRREEKQSNSTACQDL-----YAWTDGSISQFRNWYVDEPSCG--SEVCVVMYHQPS 129
              |||              |||     : |:|||   : ||   ||: |   | ||:|
Db        204 IGL-------------QDLNMEGEFVWSDGSPVGYSNWNPGEPNNGGQGEDCVMM----- 245

Qy        130 APAGIGGPYMFQWNDDRC-NMKNNFICK 156
                | |     ||||  | :  : ::|:
Db        246 --RGSG-----QWNDAFCRSYLDAWVCE 266


RESULT 31
US-08-365-103B-6
; Sequence 6, Application US/08365103B
; Patent No. 5766943
;  GENERAL INFORMATION:
;    APPLICANT:  Lynch, Richard G
;    APPLICANT:  Nunez, Raphael D.
;    APPLICANT:  Yodoi, Jungi
;    TITLE OF INVENTION:  DNA Sequences for Soluble Froms of CD23
;    TITLE OF INVENTION:  and Methods of Use for Same
;    NUMBER OF SEQUENCES:  14
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Zarley, McKee, Thomte, Voorhees & Sease
;      STREET:  801 Grand Ave. Suite 3200
;      CITY:  Des Moines
;      STATE:  Iowa
;      COUNTRY:  United States
;      ZIP:  50309
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/365,103B
;      FILING DATE:  28-DEC-1994
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Nebel, Heidi S.
;      REGISTRATION NUMBER:  37,719
;      REFERENCE/DOCKET NUMBER:  Uirf N5-24
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (515) 288-3667
;      TELEFAX:  (515) 288-1338
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  300 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-365-103B-6

  Query Match            12.7%;  Score 141.5;  DB 1;  Length 300;
  Best Local Similarity  31.8%;  Pred. No. 1.1e-06;
  Matches   47;  Conservative   21;  Mismatches   37;  Indels   43;  Gaps    9;

Qy         17 QRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFW 76
              |: ||   ||   |::  : :|: ||    |:|||| |: ||  : : |     :  | |
Db        167 QQKCY---YFGKGSKQ--WIQARFACSDLQGRLVSIHSQKEQDFLMQHI-----NKKDSW 216

Qy         77 IGLRRREEKQSNSTACQDL-----YAWTDGSISQFRNWYVDEPSCG--SEVCVVMYHQPS 129
              |||              |||     : |:|||   : ||   ||: |   | ||:|
Db        217 IGL-------------QDLNMEGEFVWSDGSPVGYSNWNPGEPNNGGQGEDCVMM----- 258

Qy        130 APAGIGGPYMFQWNDDRC-NMKNNFICK 156
                | |     ||||  | :  : ::|:
Db        259 --RGSG-----QWNDAFCRSYLDAWVCE 279


RESULT 32
US-08-365-103B-2
; Sequence 2, Application US/08365103B
; Patent No. 5766943
;  GENERAL INFORMATION:
```

```
;    APPLICANT:  Lynch, Richard G
;    APPLICANT:  Nunez, Raphael D.
;    APPLICANT:  Yodoi, Jungi
;    TITLE OF INVENTION:  DNA Sequences for Soluble Froms of CD23
;    TITLE OF INVENTION:  and Methods of Use for Same
;    NUMBER OF SEQUENCES:  14
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Zarley, McKee, Thomte, Voorhees & Sease
;      STREET:  801 Grand Ave. Suite 3200
;      CITY:  Des Moines
;      STATE:  Iowa
;      COUNTRY:  United States
;      ZIP:  50309
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/365,103B
;      FILING DATE:  28-DEC-1994
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Nebel, Heidi S.
;      REGISTRATION NUMBER:  37,719
;      REFERENCE/DOCKET NUMBER:  Uirf N5-24
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (515) 288-3667
;      TELEFAX:  (515) 288-1338
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  327 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-365-103B-2

  Query Match             12.7%;  Score 141.5;  DB 1;  Length 327;
  Best Local Similarity   31.8%;  Pred. No. 1.3e-06;
  Matches   47;  Conservative   21;  Mismatches   37;  Indels   43;  Gaps    9;

Qy         17 QRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFW 76
              |: ||   ||   |::  : :|: ||    |:|||| |: ||  : : |     :  | |
Db        194 QQKCY---YFGKGSKQ--WIQARFACSDLQGRLVSIHSQKEQDFLMQHI-----NKKDSW 243

Qy         77 IGLRRREEKQSNSTACQDL-----YAWTDGSISQFRNWYVDEPSCG--SEVCVVMYHQPS 129
              |||               |||      : |:|||   : ||   ||: |   | ||:|
Db        244 IGL-------------QDLNMEGEPVWSDGSPVGYSNWNPGEPNNGGQGEDCVMM----- 285

Qy        130 APAGIGGPYMFQWNDDRC-NMKNNFICK 156
                | |     ||||  | :  : ::|:
Db        286 --RGSG-----QWNDAFCRSYLDAWVCE 306


RESULT 33
5514582-7
;Patent No. 5514582
;    APPLICANT: CAPON, DANIEL J.;LASKY, LAURENCE A.
;    TITLE OF INVENTION: RECOMBINANT DNA ENCODING HYBRID
;IMMUNOGLOBULINS
;    NUMBER OF SEQUENCES: 43
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/185,670
;      FILING DATE: 21-JAN-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER: 986,931
;      FILING DATE: 08-DEC-1992
;      APPLICATION NUMBER: 808,122
;      FILING DATE: 16-DEC-1991
;      APPLICATION NUMBER: 440,625
;      FILING DATE: 22-NOV-1989
;      APPLICATION NUMBER: 315,015
;      FILING DATE: 23-FEB-1989
;SEQ ID NO:7:
;      LENGTH: 117
5514582-7

  Query Match             12.4%;  Score 138.5;  DB 6;  Length 117;
  Best Local Similarity   27.2%;  Pred. No. 6.4e-07;
  Matches   37;  Conservative   34;  Mismatches   38;  Indels   27;  Gaps    8;

Qy         26 FHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLP-SDGDFWIGLRRREE 84
              :| : : :|:| |:: |:::   ||:|::: |   || ::|| || |   :|||:|:
Db          3 YHYSEKPMNWENARKFCKQNYTDLVAIQNKRE---IE-YLENTLPKSPYYYWIGIRK--- 55

Qy         85 KQSNSTACQDLYAW--TDGSIS-QFRNWYVDEPS--CGSEVCVVMYHQPSAPAGIGGPYM 139
                        :: |  |: ::: :  ||   ||:     | || :| :    :|
Db         56 -------IGKMWTWVGTNKTLTKEAENWGAGEPNNKKSKEDCVEIYIKRERDSG------ 102

Qy        140 FQWNDDRCNMKNNFIC 155
               :|||| |: :   :|
Db        103 -KWNDDACHKRKAALC 117


RESULT 34
US-08-340-539A-12
; Sequence 12, Application US/08340539A
; Patent No. 5808025
;  GENERAL INFORMATION:
;    APPLICANT:  Tedder, Thomas F.
;    APPLICANT:  Kansas, Geoffrey S.
;    TITLE OF INVENTION:  CHIMERIC SELECTINS AS SIMULTANEOUS
;    TITLE OF INVENTION:  BLOCKING AGENTS FOR COMPONENT SELECTIN FUNCTION
;    NUMBER OF SEQUENCES:  28
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  FISH & NEAVE
;      STREET:  1251 Avenue of the Americas
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10020
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/340,539A
;      FILING DATE:  16-NOV-1994
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/008,459
;      FILING DATE:  25-JAN-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Gunnison, Jane
;      REGISTRATION NUMBER:  38,479
;      REFERENCE/DOCKET NUMBER:  CG-104 CON
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  212-596-9000
;      TELEFAX:  212-596-9090
;  INFORMATION FOR SEQ ID NO:  12:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  119 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-340-539A-12

  Query Match             12.4%;  Score 138.5;  DB 1;  Length 119;
  Best Local Similarity   27.2%;  Pred. No. 6.5e-07;
  Matches   37;  Conservative   34;  Mismatches   38;  Indels   27;  Gaps    8;

Qy         26 FHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLP-SDGDFWIGLRRREE 84
              :| : : :|:| |:: |:::   ||:|::: |   || ::|| || |   :|||:|:
Db          3 YHYSEKPMNWENARKFCKQNYTNLVAIQNKRE---IE-YLENTLPKSPYYYWIGIRK--- 55
```

```
Qy          85 KQSNSTACQDLYAW--TDGSIS-QFRNWYVDEPS--CGSEVCVVMYHQPSAPAGIGGPYM 139
                  :: |  |: ::: :  ||   ||:     | || :| :    :|
Db          56 -------IGKMWTWVGTNKTLTKEAENWGAGEPNNKKSKEDCVEIYIKRERDSG------ 102

Qy         140 FQWNDDRCNMKNNFIC 155
               :|||| |: :   :|
Db         103 -KWNDDACHKRKAALC 117


RESULT 35
5514582-12
;Patent No. 5514582
;     APPLICANT: CAPON, DANIEL J.;LASKY, LAURENCE A.
;     TITLE OF INVENTION: RECOMBINANT DNA ENCODING HYBRID
;IMMUNOGLOBULINS
;     NUMBER OF SEQUENCES: 43
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/185,670
;       FILING DATE: 21-JAN-1994
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER: 986,931
;       FILING DATE: 08-DEC-1992
;       APPLICATION NUMBER: 808,122
;       FILING DATE: 16-DEC-1991
;       APPLICATION NUMBER: 440,625
;       FILING DATE: 22-NOV-1989
;       APPLICATION NUMBER: 315,015
;       FILING DATE: 23-FEB-1989
;SEQ ID NO:12:
;       LENGTH: 110
5514582-12

  Query Match           12.3%;  Score 137.5;  DB 6;  Length 110;
  Best Local Similarity  32.4%;  Pred. No. 7.5e-07;
  Matches   45;  Conservative   16;  Mismatches   45;  Indels   33;  Gaps    8;

Qy          20 CYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFWIGL 79
               ||   ||   :::  :  |: ||    |||||| | :||  : |       |    ||||
Db           2 CY---YFGKGTKQ--WVHARYACDDMEGQLVSIHSPEEQDFLTKH-----ASHTGSWIGL 51

Qy          80 RRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEPSCGS--EVCVVMYHQPSAPAGIGGP 137
               |  : |          : | |||   : ||   ||:  |  | ||:|       | |
Db          52 RNLDLKGE--------FIWVDGSHVDYSNWAPGEPTSRSQGEDCVMM-------RGSG-- 94

Qy         138 YMFQWNDDRCNMK-NNFIC 155
                :|||  |: |   ::|
Db          95 ---RWNDAFCDRKLGAWVC 110


RESULT 36
5514582-19
;Patent No. 5514582
;     APPLICANT: CAPON, DANIEL J.;LASKY, LAURENCE A.
;     TITLE OF INVENTION: RECOMBINANT DNA ENCODING HYBRID
;IMMUNOGLOBULINS
;     NUMBER OF SEQUENCES: 43
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/185,670
;       FILING DATE: 21-JAN-1994
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER: 986,931
;       FILING DATE: 08-DEC-1992
;       APPLICATION NUMBER: 808,122
;       FILING DATE: 16-DEC-1991
;       APPLICATION NUMBER: 440,625
;       FILING DATE: 22-NOV-1989
;       APPLICATION NUMBER: 315,015
;       FILING DATE: 23-FEB-1989
;SEQ ID NO:19:
;       LENGTH: 123
5514582-19

  Query Match           12.3%;  Score 137.5;  DB 6;  Length 123;
  Best Local Similarity  26.9%;  Pred. No. 8.7e-07;
  Matches   36;  Conservative   26;  Mismatches   57;  Indels   15;  Gaps    5;

Qy          25 YFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFWIGLRRREE 84
               |  :|  : |: :|    | |   |||:||| |:   |   :: ::    : |:|     :
Db           2 YLIETELKYNWHQAWHECARHDQQLVTIESADKNNAIIDLVKRVVGKSHNLWLG---GND 58

Qy          85 KQSNSTACQDLYAWT-DGSISQFRNWYVDEPS--CGSEVCVVMYHQPSAPAGIGGPYMFQ 141
               : |:|      : |:  |    |  |  : |      | || ::  :  |          ::|
Db          59 EYSSSRDYGRPFFWSPTGQAFSFAYWSENNPDNYKHQEHCVHIW--DTKP-------LYQ 109

Qy         142 WNDDRCNMKNNFIC 155
               |||: ||:|  :||
Db         110 WNDNDCNVKMGYIC 123


RESULT 37
US-09-724-864-39
; Sequence 39, Application US/09724864
; Patent No. 6380362
; GENERAL INFORMATION:
;  APPLICANT:  Watson, James D
;  APPLICANT:  Murison, James G.
;  TITLE OF INVENTION:  Polynucleotides, polypeptides expressed
;  TITLE OF INVENTION:  by the polynucleotides and methods for their use.
;  FILE REFERENCE:  11000.1050U1
;  CURRENT APPLICATION NUMBER: US/09/724,864
;  CURRENT FILING DATE:  2000-11-28
;  PRIOR APPLICATION NUMBER:  U.S. No. 6380362 60/171,678
;  PRIOR FILING DATE:  1999-12-23
;  NUMBER OF SEQ ID NOS:  72
;  SOFTWARE:  FastSEQ for Windows Version 4.0
; SEQ ID NO 39
;   LENGTH: 492
;   TYPE: PRT
;   ORGANISM: Mouse
US-09-724-864-39

  Query Match           12.3%;  Score 137.5;  DB 4;  Length 492;
  Best Local Similarity  25.1%;  Pred. No. 5.7e-06;
  Matches   43;  Conservative   29;  Mismatches   66;  Indels   33;  Gaps    9;

Qy          20 CYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFWIGL 79
               || :       ||  | ||  |||  || | :  : :| : ::  :  : |::|  ||||
Db          32 CYALF-----PRRRTFLEAWRACRELGGNLATPRTPEEAQRVDSLV-GVGPANGLLWIGL 85

Qy          80 RRREEKQSNSTACQDL-----YAWTDG-SISQFRNWY--VDEPSCGSEVCVVMYHQPSAP 131
               :|:  :       ||       : || |   : | ||      |  | :: |
Db          86 QRQARQ------CQPQRPLRGFIWTTGDQDTAFTNWAQPATEGPCPAQRC---------- 129

Qy         132 AGIGGPYMFQWNDDRCNMK-NNFICKYSDEK--PAVPSREAEGEETELTTP 179
               | :       :| :  | :  : ::|::  |   ||:|    :      |||
Db         130 AALEASGEHRWLEGSCTLAVDGYLCQFGFEGACPALPLEVGQAGPAVYTTP 180


RESULT 38
US-09-111-470-8
; Sequence 8, Application US/09111470
; Patent No. 6277959
;  GENERAL INFORMATION:
;    APPLICANT:  Valladeau, Jenny
;    APPLICANT:  Ravel, Odile
;    APPLICANT:  Bates, Elizabeth E.M.
;    APPLICANT:  Ford, John
;    APPLICANT:  Saeland, Sem
;    APPLICANT:  Lebecque, Serge J.E.
;    TITLE OF INVENTION:  Mammalian Membrane Protein Genes;
;    TITLE OF INVENTION:  Related Reagents
;    NUMBER OF SEQUENCES:  11
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  DNAX Research Institute
;      STREET:  901 California Avenue
```

```
;      CITY:  Palo Alto
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94304-1104
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/111,470
;      FILING DATE:  08-JUL-1998
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/053,080
;      FILING DATE:  09-JUL-1997
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Ching, Edwin P.
;      REGISTRATION NUMBER:  34,090
;      REFERENCE/DOCKET NUMBER:  SF0695
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (650)852-9196
;      TELEFAX:  (650)496-1200
;  INFORMATION FOR SEQ ID NO:  8:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  238 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-09-111-470-8

  Query Match             12.1%;  Score 135;  DB 3;  Length 238;
  Best Local Similarity   26.6%;  Pred. No. 3.9e-06;
  Matches   37;  Conservative   24;  Mismatches   52;  Indels   26;  Gaps    5;

Qy         20 CYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFWIGL 79
Db        118 CYLV---PTVSSSASWNKSEENCSRMGAHLVVIQSQEEQ----DFITGILDTHAAYFIGL 170

Qy         80 RRREEKQSNSTACQDLYAWTDGS--ISQFRNWYVDEPSCGSEVCVVMYHQPSAPAGIGGP 137
Db        171 WDTGHRQ---------WQWVDQTPYEESITFWHNGEPSSGNEKCATIIYRWKTGWG---- 217

Qy        138 YMFQWNDDRCNMKNNFICK 156
Db        218 ----WNDISCSLKQKSVCQ 232


RESULT 39
US-09-535-521-8
; Sequence 8, Application US/09535521
; Patent No. 6410714
; GENERAL INFORMATION:
;  APPLICANT: Weber, Eric R.
;  APPLICANT: McCall, Catherine A.
;  TITLE OF INVENTION: NOVEL CANINE LOW AFFINITY IgE RECEPTOR (CANINE CD23)
;  TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES AND USES THEREOF
;  FILE REFERENCE: AL-5
;  CURRENT APPLICATION NUMBER: US/09/535,521
;  CURRENT FILING DATE: 2000-03-24
;  EARLIER APPLICATION NUMBER: 60/125,913
;  EARLIER FILING DATE: 1999-03-24
;  NUMBER OF SEQ ID NOS: 26
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
;   LENGTH: 128
;   TYPE: PRT
;   ORGANISM: Canis familiaris
US-09-535-521-8

  Query Match             12.1%;  Score 134.5;  DB 4;  Length 128;
  Best Local Similarity   28.2%;  Pred. No. 1.9e-06;
  Matches   40;  Conservative   24;  Mismatches   45;  Indels   33;  Gaps    8;

Qy         17 QRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFW 76
Db         11 QRKCY---YFGEEPKK--WIQARFACSKLQGRLASIHSQEEQDFLARY-----ANKKGTW 60

Qy         77 IGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEPSCG--SEVCVVMYHQPSAPAGI 134
Db         61 IGLRDLDREGE--------FIWMDENPLNYSNWRPGEPNNGGQGEDCVMM-------QGS 105

Qy        135 GGPYMFQWNDDRCNMK-NNFIC 155
Db        106 G-----QWNDAFCGSSLDGWVC 122


RESULT 40
US-09-535-521-11
; Sequence 11, Application US/09535521
; Patent No. 6410714
; GENERAL INFORMATION:
;  APPLICANT: Weber, Eric R.
;  APPLICANT: McCall, Catherine A.
;  TITLE OF INVENTION: NOVEL CANINE LOW AFFINITY IgE RECEPTOR (CANINE CD23)
;  TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES AND USES THEREOF
;  FILE REFERENCE: AL-5
;  CURRENT APPLICATION NUMBER: US/09/535,521
;  CURRENT FILING DATE: 2000-03-24
;  EARLIER APPLICATION NUMBER: 60/125,913
;  EARLIER FILING DATE: 1999-03-24
;  NUMBER OF SEQ ID NOS: 26
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
;   LENGTH: 139
;   TYPE: PRT
;   ORGANISM: Canis familiaris
US-09-535-521-11

  Query Match             12.1%;  Score 134.5;  DB 4;  Length 139;
  Best Local Similarity   28.2%;  Pred. No. 2.1e-06;
  Matches   40;  Conservative   24;  Mismatches   45;  Indels   33;  Gaps    8;

Qy         17 QRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFW 76
Db         22 QRKCY---YFGEEPKK--WIQARFACSKLQGRLASIHSQEEQDFLARY-----ANKKGTW 71

Qy         77 IGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEPSCG--SEVCVVMYHQPSAPAGI 134
Db         72 IGLRDLDREGE--------FIWMDENPLNYSNWRPGEPNNGGQGEDCVMM-------QGS 116

Qy        135 GGPYMFQWNDDRCNMK-NNFIC 155
Db        117 G-----QWNDAFCGSSLDGWVC 133


RESULT 41
US-09-535-521-14
; Sequence 14, Application US/09535521
; Patent No. 6410714
; GENERAL INFORMATION:
;  APPLICANT: Weber, Eric R.
;  APPLICANT: McCall, Catherine A.
;  TITLE OF INVENTION: NOVEL CANINE LOW AFFINITY IgE RECEPTOR (CANINE CD23)
;  TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES AND USES THEREOF
;  FILE REFERENCE: AL-5
;  CURRENT APPLICATION NUMBER: US/09/535,521
;  CURRENT FILING DATE: 2000-03-24
;  EARLIER APPLICATION NUMBER: 60/125,913
;  EARLIER FILING DATE: 1999-03-24
;  NUMBER OF SEQ ID NOS: 26
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 14
;   LENGTH: 141
;   TYPE: PRT
;   ORGANISM: Canis familiaris
```

US-09-535-521-14

Query Match 12.1%; Score 134.5; DB 4; Length 141;
Best Local Similarity 28.2%; Pred. No. 2.2e-06;
Matches 40; Conservative 24; Mismatches 45; Indels 33; Gaps 8;

```
Qy         17 QRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFW 76
              || ||   || :  ::  : :|: || :  |:| || |::||  : ::      :    |
Db         24 QRKCY---YFGEEPKK--WIQARFACSKLQGRLASIHSQEEQDFLARY-----ANKKGTW 73

Qy         77 IGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEPSCG--SEVCVVMYHQPSAPAGI 134
              ||||  : :          : | | :   : ||   ||: |   | ||:|       |
Db         74 IGLRDLDREGE--------FIWMDENPLNYSNWRPGEPNNGGQGEDCVMM-------QGS 118

Qy        135 GGPYMFQWNDDRCNMK-NNFIC 155
              |     ||||  |    : ::|
Db        119 G-----QWNDAFCGSSLDGWVC 135
```

RESULT 42
US-09-535-521-17
; Sequence 17, Application US/09535521
; Patent No. 6410714
; GENERAL INFORMATION:
; APPLICANT: Weber, Eric R.
; APPLICANT: McCall, Catherine A.
; TITLE OF INVENTION: NOVEL CANINE LOW AFFINITY IgE RECEPTOR (CANINE CD23)
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES AND USES THEREOF
; FILE REFERENCE: AL-5
; CURRENT APPLICATION NUMBER: US/09/535,521
; CURRENT FILING DATE: 2000-03-24
; EARLIER APPLICATION NUMBER: 60/125,913
; EARLIER FILING DATE: 1999-03-24
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 17
; LENGTH: 187
; TYPE: PRT
; ORGANISM: Canis familiaris
US-09-535-521-17

Query Match 12.1%; Score 134.5; DB 4; Length 187;
Best Local Similarity 28.2%; Pred. No. 3.2e-06;
Matches 40; Conservative 24; Mismatches 45; Indels 33; Gaps 8;

```
Qy         17 QRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFW 76
              || ||   || :  ::  : :|: || :  |:| || |::||  : ::      :    |
Db         70 QRKCY---YFGEEPKK--WIQARFACSKLQGRLASIHSQEEQDFLARY-----ANKKGTW 119

Qy         77 IGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEPSCG--SEVCVVMYHQPSAPAGI 134
              ||||  : :          : | | :   : ||   ||: |   | ||:|       |
Db        120 IGLRDLDREGE--------FIWMDENPLNYSNWRPGEPNNGGQGEDCVMM-------QGS 164

Qy        135 GGPYMFQWNDDRCNMK-NNFIC 155
              |     ||||  |    : ::|
Db        165 G-----QWNDAFCGSSLDGWVC 181
```

RESULT 43
US-09-535-521-20
; Sequence 20, Application US/09535521
; Patent No. 6410714
; GENERAL INFORMATION:
; APPLICANT: Weber, Eric R.
; APPLICANT: McCall, Catherine A.
; TITLE OF INVENTION: NOVEL CANINE LOW AFFINITY IgE RECEPTOR (CANINE CD23)
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES AND USES THEREOF
; FILE REFERENCE: AL-5
; CURRENT APPLICATION NUMBER: US/09/535,521
; CURRENT FILING DATE: 2000-03-24
; EARLIER APPLICATION NUMBER: 60/125,913
; EARLIER FILING DATE: 1999-03-24
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 20
; LENGTH: 208
; TYPE: PRT
; ORGANISM: Canis familiaris
US-09-535-521-20

Query Match 12.1%; Score 134.5; DB 4; Length 208;
Best Local Similarity 28.2%; Pred. No. 3.7e-06;
Matches 40; Conservative 24; Mismatches 45; Indels 33; Gaps 8;

```
Qy         17 QRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFW 76
              || ||   || :  ::  : :|: || :  |:| || |::||  : ::      :    |
Db         91 QRKCY---YFGEEPKK--WIQARFACSKLQGRLASIHSQEEQDFLARY-----ANKKGTW 140

Qy         77 IGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEPSCG--SEVCVVMYHQPSAPAGI 134
              ||||  : :          : | | :   : ||   ||: |   | ||:|       |
Db        141 IGLRDLDREGE--------FIWMDENPLNYSNWRPGEPNNGGQGEDCVMM-------QGS 185

Qy        135 GGPYMFQWNDDRCNMK-NNFIC 155
              |     ||||  |    : ::|
Db        186 G-----QWNDAFCGSSLDGWVC 202
```

RESULT 44
US-09-535-521-2
; Sequence 2, Application US/09535521
; Patent No. 6410714
; GENERAL INFORMATION:
; APPLICANT: Weber, Eric R.
; APPLICANT: McCall, Catherine A.
; TITLE OF INVENTION: NOVEL CANINE LOW AFFINITY IgE RECEPTOR (CANINE CD23)
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES AND USES THEREOF
; FILE REFERENCE: AL-5
; CURRENT APPLICATION NUMBER: US/09/535,521
; CURRENT FILING DATE: 2000-03-24
; EARLIER APPLICATION NUMBER: 60/125,913
; EARLIER FILING DATE: 1999-03-24
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 292
; TYPE: PRT
; ORGANISM: Canis familiaris
US-09-535-521-2

Query Match 12.1%; Score 134.5; DB 4; Length 292;
Best Local Similarity 28.2%; Pred. No. 5.8e-06;
Matches 40; Conservative 24; Mismatches 45; Indels 33; Gaps 8;

```
Qy         17 QRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFW 76
              || ||   || :  ::  : :|: || :  |:| || |::||  : ::      :    |
Db        175 QRKCY---YFGEEPKK--WIQARFACSKLQGRLASIHSQEEQDFLARY-----ANKKGTW 224

Qy         77 IGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEPSCG--SEVCVVMYHQPSAPAGI 134
              ||||  : :          : | | :   : ||   ||: |   | ||:|       |
Db        225 IGLRDLDREGE--------FIWMDENPLNYSNWRPGEPNNGGQGEDCVMM-------QGS 269

Qy        135 GGPYMFQWNDDRCNMK-NNFIC 155
              |     ||||  |    : ::|
Db        270 G-----QWNDAFCGSSLDGWVC 286
```

RESULT 45
US-09-535-521-5
; Sequence 5, Application US/09535521
; Patent No. 6410714
; GENERAL INFORMATION:
; APPLICANT: Weber, Eric R.
; APPLICANT: McCall, Catherine A.
; TITLE OF INVENTION: NOVEL CANINE LOW AFFINITY IgE RECEPTOR (CANINE CD23)
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES AND USES THEREOF
; FILE REFERENCE: AL-5

```
;  CURRENT APPLICATION NUMBER: US/09/535,521
;  CURRENT FILING DATE: 2000-03-24
;  EARLIER APPLICATION NUMBER: 60/125,913
;  EARLIER FILING DATE: 1999-03-24
;  NUMBER OF SEQ ID NOS: 26
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
;   LENGTH: 292
;   TYPE: PRT
;   ORGANISM: Canis familiaris
US-09-535-521-5

  Query Match             12.1%;  Score 134.5;  DB 4;  Length 292;
  Best Local Similarity   28.2%;  Pred. No. 5.8e-06;
  Matches   40;  Conservative   24;  Mismatches   45;  Indels   33;  Gaps    8;

Qy         17 QRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFW 76
              || ||    || :  ::  : :|: || :  |:| || |::||  : ::      :    |
Db        175 QRKCY---YFGEEPKK--WIQARFACSKLQGRLASIHSQEEQDFLARY-----ANKKGTW 224

Qy         77 IGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEPSCG--SEVCVVMYHQPSAPAGI 134
              ||||  : :          : | | :   : ||   ||: |   | ||:|        |
Db        225 IGLRDLDREGE--------FIWMDENPLNYSNWRPGEPNNGGQGEDCVMM-------QGS 269

Qy        135 GGPYMFQWNDDRCNMK-NNFIC 155
              |     ||||  |    : ::|
Db        270 G-----QWNDAFCGSSLDGWVC 286


RESULT 46
5514582-9
;Patent No. 5514582
;     APPLICANT: CAPON, DANIEL J.;LASKY, LAURENCE A.
;     TITLE OF INVENTION: RECOMBINANT DNA ENCODING HYBRID
;IMMUNOGLOBULINS
;     NUMBER OF SEQUENCES: 43
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/185,670
;       FILING DATE: 21-JAN-1994
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER: 986,931
;       FILING DATE: 08-DEC-1992
;       APPLICATION NUMBER: 808,122
;       FILING DATE: 16-DEC-1991
;       APPLICATION NUMBER: 440,625
;       FILING DATE: 22-NOV-1989
;       APPLICATION NUMBER: 315,015
;       FILING DATE: 23-FEB-1989
;SEQ ID NO:9:
;       LENGTH: 110
5514582-9

  Query Match             12.0%;  Score 133.5;  DB 6;  Length 110;
  Best Local Similarity   26.5%;  Pred. No. 2e-06;
  Matches   35;  Conservative   20;  Mismatches   42;  Indels   35;  Gaps    6;

Qy         32 RLNFEEAKEACR--RDGGQLVSIESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNS 89
              :: : :|: ||:     |  | :|:|: |    | : : |         ||||         |
Db          6 QVRWNDAQLACQTVHPGAYLATIQSQLENAFISETVSN-----NRLWIGL---------ND 52

Qy         90 TACQDLYAWTDGSISQFRNWYVDEPS------CGSEVCVVMYHQPSAPAGIGGPYMFQWN 143
                 :  | |::|  : |  |  : |:       ||    || |   :              ||:
Db         53 IDLEGHYVWSNGEATDFTYWSSNNPNNWENQDCG----VVNYDTVTG----------QWD 98

Qy        144 DDRCNMKNNFIC 155
              || ||    ||:|
Db         99 DDDCNKNKNFLC 110


RESULT 47
US-08-340-539A-13
; Sequence 13, Application US/08340539A
; Patent No. 5808025
;  GENERAL INFORMATION:
;    APPLICANT:  Tedder, Thomas F.
;    APPLICANT:  Kansas, Geoffrey S.
;    TITLE OF INVENTION:  CHIMERIC SELECTINS AS SIMULTANEOUS
;    TITLE OF INVENTION:  BLOCKING AGENTS FOR COMPONENT SELECTIN FUNCTION
;    NUMBER OF SEQUENCES:  28
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  FISH & NEAVE
;      STREET:  1251 Avenue of the Americas
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10020
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/340,539A
;      FILING DATE:  16-NOV-1994
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/008,459
;      FILING DATE:  25-JAN-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Gunnison, Jane
;      REGISTRATION NUMBER:  38,479
;      REFERENCE/DOCKET NUMBER:  CG-104 CON
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  212-596-9000
;      TELEFAX:  212-596-9090
;  INFORMATION FOR SEQ ID NO:  13:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  119 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-340-539A-13

  Query Match             12.0%;  Score 133.5;  DB 1;  Length 119;
  Best Local Similarity   26.5%;  Pred. No. 2.2e-06;
  Matches   36;  Conservative   32;  Mismatches   41;  Indels   27;  Gaps    8;

Qy         26 FHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGD-FWIGLRRREE 84
              :| : | :|:| |:: |: :   ||:|::: |   || ::|  || :   :|||:|:
Db          3 YHYSERSMNWENARKFCKHNYTDLVAIQNKRE---IE-YLEKTLPKNPTYYWIGIRK--- 55

Qy         85 KQSNSTACQDLYAW--TDGSIS-QFRNWYVDEPS--CGSEVCVVMYHQPSAPAGIGGPYM 139
                          : |  |: ::: :  ||   ||:     | || :| :    :|
Db         56 -------IGKTWTWVGTNKTLTKEAENWGTGEPNNKKSKEDCVEIYIKRERDSG------ 102

Qy        140 FQWNDDRCNMKNNFIC 155
               :|||| |: :   :|
Db        103 -KWNDDACHKRKAALC 117


RESULT 48
US-07-893-929A-9
; Sequence 9, Application US/07893929A
; Patent No. 5336667
;  GENERAL INFORMATION:
;    APPLICANT:  Kirby, Edward P.
;    APPLICANT:  Peng, Man-ling
;    TITLE OF INVENTION:  Alboaggregins: Platelet
;    TITLE OF INVENTION:  Agonists Which Bind To Platelet
;    TITLE OF INVENTION:  Membrane Glycoprotein Ib
;    NUMBER OF SEQUENCES:  10
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Temple University - Of The Common-
;      ADDRESSEE:  wealth System of Higher Education
;      STREET:  406 University Services Building
```

```
;       CITY:  Philadelphia
;       STATE:  Pennsylvania
;       COUNTRY:  U.S.A.
;       ZIP:  19122
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Diskette, 3.50 inch, 720 Kb
;       COMPUTER:  IBM PS/2
;       OPERATING SYSTEM:  MS-DOS
;       SOFTWARE:  WordPerfect 5.1
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/07/893,929A
;       FILING DATE:  19920605
;       CLASSIFICATION:  435
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  07/803,630
;       FILING DATE:  December 3, 1991
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Monaco, Daniel A.
;       REGISTRATION NUMBER:  30,480
;       REFERENCE/DOCKET NUMBER:  6056-126 (CIP) 1
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (215) 568-8383
;       TELEFAX:  (215) 568-5549
;       TELEX:  No. 5336667e
;   INFORMATION FOR SEQ ID NO:  9:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  133 amino acids
;       TYPE:  AMINO ACID
;       TOPOLOGY:  linear
US-07-893-929A-9

  Query Match             12.0%;  Score 133.5;  DB 1;  Length 133;
  Best Local Similarity   29.1%;  Pred. No. 2.5e-06;
  Matches   41;  Conservative   21;  Mismatches   48;  Indels   31;  Gaps    8;

Qy         20 CYKVIYFHDTSRRLNFEEAKEACRR-DGGQLVSIESEDEQKLIEKFIENLLPS---DGDF 75
              ||::       |   ::||::  | : ||| ||||||  : : : : |   ||    :
Db         15 CYQLF------RLKTWDEAEKYCNQWDGGHLVSIESNAKAEFVAQLISRKLPKSAIEDRV 68

Qy         76 WIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEPSCGSEVCVVMYHQPSAPAGIG 135
              ||||| | :::     |  |  ||| |   :   :|  |:     | |:  |
Db         69 WIGLRDRSKREQ----CGHL--WTDNSFVHYE--HVVPPT----KCFVLERQTE------ 110

Qy        136 GPYMFQWNDDRCNMKNNFICK 156
                :|    |  |  |:||
Db        111 ---FRKWIAVNCEFKFPFVCK 128


RESULT 49
PCT-US92-10344-9
; Sequence 9, Application PC/TUS9210344
;  GENERAL INFORMATION:
;    APPLICANT:  Kirby, Edward P.
;    APPLICANT:  Peng, Man-ling
;    TITLE OF INVENTION:  Alboaggregins: Platelet
;    TITLE OF INVENTION:  Agonists Which Bind To Platelet
;    TITLE OF INVENTION:  Membrane Glycoprotein Ib
;    NUMBER OF SEQUENCES:  10
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Temple University - Of The Common-
;      ADDRESSEE:  wealth System of Higher Education
;      STREET:  406 University Services Building
;      CITY:  Philadelphia
;      STATE:  Pennsylvania
;      COUNTRY:  U.S.A.
;      ZIP:  19122
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.50 inch, 720 Kb
;      COMPUTER:  IBM PS/2
;      OPERATING SYSTEM:  MS-DOS
;      SOFTWARE:  WordPerfect 5.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US92/10344
;      FILING DATE:  19921201
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/803,630
;      FILING DATE:  December 3, 1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Monaco, Daniel A.
;      REGISTRATION NUMBER:  30,480
;      REFERENCE/DOCKET NUMBER:  6056-126 (CIP) 1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (215) 568-8383
;      TELEFAX:  (215) 568-5549
;      TELEX:  None
;  INFORMATION FOR SEQ ID NO:  9:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  133 amino acids
;      TYPE:  AMINO ACID
;      TOPOLOGY:  linear
PCT-US92-10344-9

  Query Match             12.0%;  Score 133.5;  DB 5;  Length 133;
  Best Local Similarity   29.1%;  Pred. No. 2.5e-06;
  Matches   41;  Conservative   21;  Mismatches   48;  Indels   31;  Gaps    8;

Qy         20 CYKVIYFHDTSRRLNFEEAKEACRR-DGGQLVSIESEDEQKLIEKFIENLLPS---DGDF 75
              ||::       |   ::||::  | : ||| ||||||  : : : : |   ||    :
Db         15 CYQLF------RLKTWDEAEKYCNQWDGGHLVSIESNAKAEFVAQLISRKLPKSAIEDRV 68

Qy         76 WIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEPSCGSEVCVVMYHQPSAPAGIG 135
              ||||| | :::     |  |  ||| |   :   :|  |:     . | |:  |
Db         69 WIGLRDRSKREQ----CGHL--WTDNSFVHYE--HVVPPT----KCFVLERQTE------ 110

Qy        136 GPYMFQWNDDRCNMKNNFICK 156
                :|    |  |  |:||
Db        111 ---FRKWIAVNCEFKFPFVCK 128


RESULT 50
US-09-111-470-6
; Sequence 6, Application US/09111470
; Patent No. 6277959
;  GENERAL INFORMATION:
;    APPLICANT:  Valladeau, Jenny
;    APPLICANT:  Ravel, Odile
;    APPLICANT:  Bates, Elizabeth E.M.
;    APPLICANT:  Ford, John
;    APPLICANT:  Saeland, Sem
;    APPLICANT:  Lebecque, Serge J.E.
;    TITLE OF INVENTION:  Mammalian Membrane Protein Genes;
;    TITLE OF INVENTION:  Related Reagents
;    NUMBER OF SEQUENCES:  11
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  DNAX Research Institute
;      STREET:  901 California Avenue
;      CITY:  Palo Alto
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94304-1104
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/111,470
;      FILING DATE:  08-JUL-1998
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/053,080
;      FILING DATE:  09-JUL-1997
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Ching, Edwin P.
;      REGISTRATION NUMBER:  34,090
```

```
;      REFERENCE/DOCKET NUMBER:  SF0695
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (650)852-9196
;      TELEFAX:  (650)496-1200
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  287 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  not relevant
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-09-111-470-6

  Query Match             12.0%;  Score 133.5;  DB 3;  Length 287;
  Best Local Similarity   27.4%;  Pred. No. 7.2e-06;
  Matches   48;  Conservative   23;  Mismatches   51;  Indels   53;  Gaps   10;

Qy         15 GTQRPCYKV--------IYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIE 66
              |:|| |  |         |:   | :  : ||:: |: :   || | | :||| |   ::
Db        147 GSQRTCCPVNWVEHQGSCYWFSHSGKA-WAEAEKYCQLENAHLVVINSWEEQKFI---VQ 202

Qy         67 NLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGS--ISQFRNWYVDEPS-------CG 117
              :  |   : ||||          |     : | ||:     ::|| | :|         |
Db        203 HTNPF--NTWIGL----------TDSDGSWKWVDGTDYRHNYKNWAVTQPDNWHGHELGG 250

Qy        118 SEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGE 172
              || || :  ||            :|||| |     ::|:          | | ||
Db        251 SEDCVEV--QPDG----------RWNDDFCLQVYRWVCE--------KRRNATGE 285

Search completed: September  9, 2004, 22:47:17
Job time : 34 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: September 9, 2004, 22:23:07 ; Search time 124 Seconds
(without alignments)
469.393 Million cell updates/sec

Title: US-09-887-855-2_COPY_22_227
Perfect score: 1115
Sequence: 1 ATGRLLSGQPVCRGGTQRPC.....EEDAKKTFKESREAALNLAY 206

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 50 summaries

Database : A_Geneseq_29Jan04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1115	100.0	374	3	AAY93948	Aay93948 Amino aci
2	1115	100.0	374	4	AAE03651	Aae03651 Human ext
3	1115	100.0	374	5	ABB90203	Abb90203 Human pol
4	1115	100.0	374	5	ABG66680	Abg66680 Human nov
5	1115	100.0	374	6	ADA54522	Ada54522 Human pro
6	1115	100.0	387	4	AAM25796	Aam25796 Human pro
7	1111	99.6	374	6	ADA54574	Ada54574 Human pro
8	1105	99.1	374	3	AAY91490	Aay91490 Human sec
9	1101	98.7	382	2	AAY13367	Aay13367 Amino aci
10	1101	98.7	382	3	ADC78457	Adc78457 Human PRO
11	1101	98.7	382	4	AAB80235	Aab80235 Human PRO
12	1101	98.7	382	4	AAU29033	Aau29033 Human PRO
13	1101	98.7	382	6	ABU58409	Abu58409 Human PRO
14	1101	98.7	382	6	ABU71613	Abu71613 Human PRO
15	1101	98.7	382	6	ABU87957	Abu87957 Novel hum
16	1101	98.7	382	6	ABU84272	Abu84272 Human sec
17	1101	98.7	382	6	ABR66146	Abr66146 Human sec
18	1101	98.7	382	6	ABR65536	Abr65536 Human sec
19	1101	98.7	382	6	ABU99476	Abu99476 Human sec
20	1101	98.7	382	6	ABU82715	Abu82715 Human PRO
21	1101	98.7	382	6	ABU89836	Abu89836 Novel hum
22	1101	98.7	382	6	ABU71468	Abu71468 Human PRO
23	1101	98.7	382	6	ABR68085	Abr68085 Human sec
24	1101	98.7	382	6	ABU96138	Abu96138 Novel hum
25	1101	98.7	382	6	ABU92569	Abu92569 Human sec
26	1101	98.7	382	6	ABO08646	Abo08646 Human sec
27	1101	98.7	382	6	ABO02698	Abo02698 Human sec
28	1101	98.7	382	6	ABR74852	Abr74852 Human sec
29	1101	98.7	382	6	ABR94614	Abr94614 Human sec
30	1101	98.7	382	6	ABU85587	Abu85587 Human PRO
31	1101	98.7	382	6	ABU98747	Abu98747 Novel hum
32	1101	98.7	382	6	ABU97962	Abu97962 Novel hum
33	1101	98.7	382	6	ABU91668	Abu91668 Novel hum
34	1101	98.7	382	6	ABU71914	Abu71914 Human sec
35	1101	98.7	382	6	ABU89361	Abu89361 Human PRO
36	1101	98.7	382	6	ABU86202	Abu86202 Human sec
37	1101	98.7	382	6	ABU67415	Abu67415 Human sec
38	1101	98.7	382	6	ABU80443	Abu80443 Human PRO
39	1101	98.7	382	6	ABO01797	Abo01797 Novel hum
40	1101	98.7	382	6	ABR99361	Abr99361 Human sec
41	1101	98.7	382	6	ABR98751	Abr98751 Human sec
42	1101	98.7	382	6	ABO16274	Abo16274 Human sec
43	1101	98.7	382	6	ABR92174	Abr92174 Human sec
44	1101	98.7	382	6	ABO18815	Abo18815 Human sec
45	1101	98.7	382	6	ABR78236	Abr78236 Human sec
46	1101	98.7	382	6	ABU84972	Abu84972 Novel hum
47	1101	98.7	382	6	ABO00111	Abo00111 Novel hum
48	1101	98.7	382	6	ABO11443	Abo11443 Human sec
49	1101	98.7	382	6	ABO02088	Abo02088 Human sec
50	1101	98.7	382	6	ABU54370	Abu54370 Human sec

ALIGNMENTS

RESULT 1
AAY93948
ID AAY93948 standard; protein; 374 AA.
XX
AC AAY93948;
XX
DT 03-OCT-2000 (first entry)
XX
DE Amino acid sequence of a lectin ss3939 polypeptide.
XX
KW Human; lectin ss3939; chromosome 11; gene therapy.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Peptide 1. .21
FT /note= "signal peptide"
FT Domain 22. .227
FT /note= "extracellular coding region"
FT Domain 228. .248
FT /note= "predicted transmembrane domain"
FT Domain 249. .374
FT /note= "predicted cytoplasmic or intracellular domain"
XX
PN WO200039296-A1.
XX
PD 06-JUL-2000.
XX
PF 22-DEC-1999; 99WO-US030523.
XX
PR 23-DEC-1998; 98US-0113820P.
XX
PA (IMMV) IMMUNEX CORP.
XX
PI Anderson DA;
XX
DR WPI; 2000-452394/39.
DR N-PSDB; AAA57382.
XX
PT ss3939 nucleic acids, polypeptides and antibodies, useful for identifying
PT human chromosome 11 and diseases associated with it.
XX
PS Claim 12; Page 8; 73pp; English.

XX
CC The present sequence represents a human lectin ss3939 polypeptide. The
CC polynucleotide sequence is a source of probes, which may be used to
CC identify nucleic acids encoding ss3939 proteins, to identify human
CC chromosome number 11, to map genes on human chromosome number 11, to
CC identify diseases associated with chromosome 11, as single-stranded sense
CC or antisense oligonucleotides to inhibit expression of polypeptides
CC encoded by the ss3939 gene, and for gene therapy. The ss3939 polypeptides
CC may be useful for developing treatments for diseases (none specified)
CC associated with defective or insufficient amounts of the polypeptides.
CC The antibodies may be useful for detecting the presence of ss3939
CC polypeptides
XX
SQ Sequence 374 AA;

Query Match 100.0%; Score 1115; DB 3; Length 374;
Best Local Similarity 100.0%; Pred. No. 1.9e-104;
Matches 206; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 ATGRLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKL 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ATGRLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKL 81

Qy         61 IEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEPSCGSEV 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 IEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEPSCGSEV 141

Qy        121 CVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGEETELTTPV 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 CVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGEETELTTPV 201

Qy        181 LPEETQEEDAKKTFKESREAALNLAY 206
              |||||||||||||||||||||||||
Db        202 LPEETQEEDAKKTFKESREAALNLAY 227
```

RESULT 2
AAE03651
ID AAE03651 standard; protein; 374 AA.
XX
AC AAE03651;
XX
DT 06-AUG-2001 (first entry)
XX
DE Human extracellular matrix and cell adhesion molecule-15 (XMAD-15).
XX
KW Human; extracellular matrix and cell adhesion molecule; XMAD;
KW gene therapy; genetic disorder; adrenoleukodystrophy; leukaemia;
KW Down's syndrome; cystic fibrosis; Gaucher's disease; myotonic dystrophy;
KW sickle cell anaemia; thalassaemia; autoimmune disorder; adenocarcinoma;
KW inflammatory disorder; acquired immune deficiency syndrome; AIDS;
KW Addison's disease; allergy; anaemia; asthma; atherosclerosis; melanoma;
KW Crohn's disease; diabetes mellitus; atopic dermatitis; lymphoma; cancer;
KW glomerulonephritis; multiple sclerosis; Grave's disease; osteoarthritis;
KW osteoporosis; psoriasis; rheumatoid arthritis; ulcerative colitis;
KW infection; cell proliferative disorder; actinic keratosis; myeloma;
KW arteriosclerosis; nootropic; anticonvulsant; antithyroid; nephrotropic;
KW neuroprotective; dermatological.
XX
OS Homo sapiens.
XX

FH	Key	Location/Qualifiers
FT	Peptide	1. .24
FT		/label= Signal_peptide
FT	Protein	25. .374
FT		/note= "Mature human extracellular matrix and cell
FT		adhesion molecule (XMAD)"
FT	Domain	46. .63
FT		/note= "C-type lectin domain"
FT	Domain	163. .176
FT		/note= "C-type lectin domain"
FT	Domain	224. .247
FT		/note= "Transmembrane motif"
FT	Domain	328. .348
FT		/note= "Transmembrane motif"

XX
PN WO200142285-A2.
XX
PD 14-JUN-2001.
XX
PF 05-DEC-2000; 2000WO-US032990.
XX
PR 10-DEC-1999; 99US-0172852P.
PR 16-DEC-1999; 99US-0172354P.
XX
PA (INCY-) INCYTE GENOMICS INC.
XX
PI Yue H, Tang YT, Lal P, Burford N, Azimzai Y, Patterson C;
PI Baughn MR, Lu DAM, Shah P, Au-Young J;
XX
DR WPI; 2001-381632/40.
DR N-PSDB; AAD08059.
XX
PT New human extracellular matrix and cell adhesion molecules and
PT polynucleotide sequences encoding them, useful for diagnosis, prevention,
PT treatment of genetic, autoimmune and cell proliferative disorders.
XX
PS Claim 1; Page 108-109; 135pp; English.
XX
CC The present sequence is a human extracellular matrix and cell adhesion
CC molecule (XMAD). The XMAD is used for screening a compound for
CC effectiveness as an agonist or antagonist of XMAD. The identified agonist
CC or antagoinst are used for treating a disease or condition associated
CC with decreased or increased expression of functional XMAD. The
CC polynucleotides encoding XMAD are useful in somatic or germline gene
CC therapy to correct a genetic deficiency, to express a conditionally
CC lethal gene product and to express a protein which affords protection
CC against intracellular parasites and also for diagnosis of disorders
CC associated with expression of XMAD. They are also used for generating
CC hybridisation probes useful in mapping the naturally occurring genomic
CC sequences and to create knock in humanised animals (pigs) or transgenic
CC animals (mice or rats) to model human diseases. Oligonucleotide or longer
CC fragments derived from the polynucleotide sequences may be used as
CC elements on a microarray. Antibodies which specifically bind XMAD may be
CC used for the diagnosis of disorders associated with the expression of
CC XMAD, or in assays to monitor patients being treated with XMAD. Diseases
CC diagnosed, prevented or treated include genetic disorders such as
CC adrenoleukodystrophy, Down's syndrome, cystic fibrosis, Gaucher's
CC disease, myotonic dystrophy, sickle cell anaemia, thalassaemia,
CC autoimmune/inflammatory disorders such as acquired immune deficiency
CC syndrome (AIDS), Addison's disease, allergies, anaemia, asthma,
CC atherosclerosis, Crohn's disease, diabetes mellitus, atopic dermatitis,
CC glomerulonephritis, multiple sclerosis, Grave's disease, osteoarthritis,
CC osteoporosis, psoriasis, rheumatoid arthritis, ulcerative colitis,
CC bacterial, fungal, parasitic, protozoal and helminthic infections and
CC cell proliferative disorders such as actinic keratosis, arteriosclerosis
CC and cancer including breast, bladder, bone marrow, brain and uterus
CC cancer, leukaemia, adenocarcinoma, lymphoma, melanoma and myeloma
XX
SQ Sequence 374 AA;

Query Match 100.0%; Score 1115; DB 4; Length 374;
Best Local Similarity 100.0%; Pred. No. 1.9e-104;
Matches 206; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 ATGRLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKL 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ATGRLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKL 81

Qy         61 IEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEPSCGSEV 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 IEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEPSCGSEV 141

Qy        121 CVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGEETELTTPV 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 CVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGEETELTTPV 201
```

```
Qy        181 LPEETQEEDAKKTFKESREAALNLAY 206
              |||||||||||||||||||||||||
Db        202 LPEETQEEDAKKTFKESREAALNLAY 227
```

RESULT 3
ABB90203
ID ABB90203 standard; protein; 374 AA.
XX
AC ABB90203;
XX
DT 24-MAY-2002 (first entry)
XX
DE Human polypeptide SEQ ID NO 2579.
XX
KW Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
KW antiallergic; hepatotropic; antidiabetic; antiinflammatory; antiulcer;
KW vulnerary; anticonvulsant; antibacterial; antifungal; antiparasitic;
KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
KW neurological disease; infection; human; secreted protein.
XX
OS Homo sapiens.
XX
PN WO200190304-A2.
XX
PD 29-NOV-2001.
XX
PF 18-MAY-2001; 2001WO-US016450.
XX
PR 19-MAY-2000; 2000US-0205515P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Birse CE, Rosen CA;
XX
DR WPI; 2002-122018/16.
DR N-PSDB; ABL90612.
XX
PT Novel 1405 isolated polypeptides, useful for diagnosis, treatment and
PT prevention of neural, immune system, muscular, reproductive,
PT gastrointestinal, pulmonary, cardiovascular, renal and proliferative
PT disorders.
XX
PS Claim 11; SEQ ID NO 2579; 2081pp + Sequence Listing; English.
XX
CC The invention relates to novel genes (ABL89449-ABL90853) and proteins
CC (ABB89040-ABB90444) useful for preventing, treating or ameliorating
CC medical conditions e.g. by protein or gene therapy. The genes are
CC isolated from a range of human tissues disclosed in the specification.
CC The nucleic acids, proteins, antibodies and (ant)agonists are useful in
CC the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and
CC ovarian cancer and other cancers of the adrenal gland, bone, bone marrow,
CC breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune
CC disorders e.g. Addison's disease, allergies, autoimmune haemolytic
CC anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,
CC multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c)
CC cardiovascular disorders such as myocardial ischaemias; (d) wound healing
CC ; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f)
CC infectious diseases such as viral, bacterial, fungal and parasitic
CC infections. Note: The sequence data for this patent did not form part of
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 374 AA;

Query Match 100.0%; Score 1115; DB 5; Length 374;
Best Local Similarity 100.0%; Pred. No. 1.9e-104;
Matches 206; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 ATGRLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKL 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ATGRLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKL 81
```

```
Qy         61 IEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEPSCGSEV 120
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 IEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEPSCGSEV 141

Qy        121 CVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGEETELTTPV 180
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 CVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGEETELTTPV 201

Qy        181 LPEETQEEDAKKTFKESREAALNLAY 206
              |||||||||||||||||||||||||
Db        202 LPEETQEEDAKKTFKESREAALNLAY 227
```

RESULT 4
ABG66680
ID ABG66680 standard; protein; 374 AA.
XX
AC ABG66680;
XX
DT 30-AUG-2002 (first entry)
XX
DE Human novel polypeptide #15.
XX
KW Human; inflammatory condition; shock; sepsis; immune response; cancer;
KW wound healing; central nervous system disease; haematopoiesis;
KW peripheral nervous system disease; amyotrophic lateral sclerosis; tendon;
KW myeloid cell disorder; lymphoid cell disorder; platelet disorder; bone;
KW cartilage; ligament; nerve tissue; ulcer; osteoporosis; osteoarthritis;
KW bone degenerative disorder; periodontal disease; reperfusion injury;
KW lung fibrosis; liver fibrosis; autoimmune disorder; bacterial infection;
KW allergic condition; thrombolysis; thrombosis; coagulation disorder;
KW fungal infection.
XX
OS Homo sapiens.
XX
PN WO200244340-A2.
XX
PD 06-JUN-2002.
XX
PF 30-NOV-2001; 2001WO-US047004.
XX
PR 30-NOV-2000; 2000US-00728952.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Goodrich RW, Liu C, Zhou P, Asundi V, Wang J, Wang D;
PI Yamazaki V, Ujwal ML, Drmanac RT;
XX
DR WPI; 2002-508509/54.
DR N-PSDB; ABK94904.
XX
PT Novel nucleic acids and polypeptides for diagnosis, treatment of
PT inflammatory, autoimmune, nervous system, myeloid or lymphoid cell
PT disorders, cancer and promoting wound healing.
XX
PS Claim 10; Page 579-580; 672pp; English.
XX
CC The invention relates to human novel polynucleotides and associated
CC polypeptides. The polynucleotides and polypeptides are useful for
CC treating inflammatory conditions such as arthritis, nephritis, Crohn's
CC disease, ischaemia-reperfusion injury, shock, sepsis, immune responses
CC and cancer and for promoting wound healing. The sequences are used to
CC induce the proliferation of neural cells and regeneration of nerve and
CC brain tissue, and are useful for the treatment of central and peripheral
CC nervous system diseases and neuropathies, such as Alzheimer's disease,
CC Parkinson's disease, Huntington's disease and amyotrophic lateral
CC sclerosis. The sequences are involved in chemotactic or chemokinetic
CC activity, regulation of haematopoiesis, treatment of myeloid or lymphoid
CC cell disorders and platelet disorders such as thrombocytopenia,
CC regeneration of bone, cartilage, tendon, ligament and/or nerve tissue
CC growth, tissue repair, healing of burns, incisions, ulcers, treatment of
CC osteoporosis, osteoarthritis, bone degenerative disorders and periodontal

CC disease. The sequences of the invention are also useful for gut
CC protection or regeneration and treatment of lung or liver fibrosis,
CC reperfusion injury in various tissues, immune deficiencies and disorders
CC including severe combined immunodeficiency (SCID), bacterial or fungal
CC infections, autoimmune disorders e.g. multiple sclerosis and myasthenia
CC gravis, allergic conditions such as asthma, thrombolysis or thrombosis
CC and coagulation disorders. Sequences ABG66666-ABG66758 represent human
CC novel polypeptides of the invention
XX
SQ Sequence 374 AA;

Query Match 100.0%; Score 1115; DB 5; Length 374;
Best Local Similarity 100.0%; Pred. No. 1.9e-104;
Matches 206; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 ATGRLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKL 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ATGRLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKL 81

Qy         61 IEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEPSCGSEV 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 IEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEPSCGSEV 141

Qy        121 CVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGEETELTTPV 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 CVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGEETELTTPV 201

Qy        181 LPEETQEEDAKKTFKESREAALNLAY 206
              |||||||||||||||||||||||||
Db        202 LPEETQEEDAKKTFKESREAALNLAY 227
```

RESULT 5
ADA54522
ID ADA54522 standard; protein; 374 AA.
XX
AC ADA54522;
XX
DT 20-NOV-2003 (first entry)
XX
DE Human protein, SEQ ID 2090.
XX
KW Cytostatic; Anti-inflammatory; Osteopathic; Neuroprotective; Nootropic;
KW Gene Therapy; human; secretory protein; membrane proteins; cancer;
KW inflammatory disease; osteoporosis; neurological disease.
XX
OS Homo sapiens.
XX
PN EP1293569-A2.
XX
PD 19-MAR-2003.
XX
PF 21-MAR-2002; 2002EP-00006586.
XX
PR 14-SEP-2001; 2001JP-00328381.
PR 24-JAN-2002; 2002US-0350435P.
XX
PA (HELI-) HELIX RES INST.
PA (REAS-) RES ASSOC BIOTECHNOLOGY.
XX
PI Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;
PI Yamamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Irie R, Tamechika I;
PI Seki N, Yoshikawa T, Otsuka M, Nagahari K, Masuho Y;
XX
DR WPI; 2003-395539/38.
DR N-PSDB; ADA52883.
XX
PT New polynucleotides encoding full-length polypeptides, e.g. secretory
PT and/or membrane proteins, useful for developing medicines for diseases in
PT which the gene is involved, or as target molecules for gene therapy.
XX
PS Claim 14; SEQ ID NO 2090; 205pp; English.
XX

CC The present invention relates to novel human secretory or membrane
CC proteins (ADA54072-ADA55710) and their coding sequences (ADA52433-
CC ADA54071). The coding sequences are useful in the gene therapy of
CC diseases caused by abnormalities of the proteins, e.g. cancer,
CC inflammatory diseases, osteoporosis or neurological disease.
XX
SQ Sequence 374 AA;

Query Match 100.0%; Score 1115; DB 6; Length 374;
Best Local Similarity 100.0%; Pred. No. 1.9e-104;
Matches 206; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 ATGRLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKL 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ATGRLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKL 81

Qy         61 IEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEPSCGSEV 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 IEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEPSCGSEV 141

Qy        121 CVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGEETELTTPV 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 CVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGEETELTTPV 201

Qy        181 LPEETQEEDAKKTFKESREAALNLAY 206
              |||||||||||||||||||||||||
Db        202 LPEETQEEDAKKTFKESREAALNLAY 227
```

RESULT 6
AAM25796
ID AAM25796 standard; protein; 387 AA.
XX
AC AAM25796;
XX
DT 16-OCT-2001 (first entry)
XX
DE Human protein sequence SEQ ID NO:1311.
XX
KW Human; cancer; ulcer; HIV infection; human immunodeficiency virus;
KW antiinflammatory; antirheumatic; antiarthritic; immunosuppressive;
KW antibacterial; endocrine; cardiant; central nervous system; virucide;
KW anti-HIV; fungicide; antimutagen; cardiovascular; antianaemic; anaemia;
KW antiaggregant; haemostatic; vulnerary; antiulcer; osteopathic; eczema;
KW dermatological; antiallergic; antiasthmatic; antidiabetic; cytostatic;
KW neuroprotective; antidepressant; nootropic; antiparkinsonian; infection;
KW immunostimulant; gene therapy; antisense therapy; vaccine; inflammation;
KW antianaphylactic; rheumatoid arthritis; septic shock; pancreatitis;
KW cardiac dysfunction; neuropathology; cardiac anaphylaxis; autoimmunity;
KW genetic disease; haematopoietic disorder; platelet disorder; asthma;
KW thrombocytopaenia; osteoporosis; severe combined immunodeficiency;
KW allergic rhinitis; diabetes; multiple sclerosis; depression;
KW Alzheimer's disease; Parkinson's disease; neurodegenerative disorder;
KW neurological disorder.
XX
OS Homo sapiens.
XX
PN WO200153455-A2.
XX
PD 26-JUL-2001.
XX
PF 22-DEC-2000; 2000WO-US035017.
XX
PR 23-DEC-1999; 99US-00471275.
PR 21-JAN-2000; 2000US-00488725.
PR 25-APR-2000; 2000US-00552317.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Drmanac RT;
XX
DR WPI; 2001-457603/49.
DR N-PSDB; AAH99737.

XX
PT Isolated human polynucleotides encoding polypeptides, useful for the
PT treatment and diagnosis of e.g. cancer, ulcers and HIV infection.
XX
PS Claim 20; Page 272; 1217pp; English.
XX
CC AAH99166 to AAH99904 encode the human proteins given in AAM25225 to
CC AAM25963. The proteins can have activities based on the tissues and cells
CC they are expressed in, such as: antiinflammatory; antirheumatic;
CC antiarthritic; immunosuppressive; antibacterial; endocrine; cardiant;
CC central nervous system; virucide; anti-HIV; fungicide; antimutagen;
CC cardiovascular; antianaemic; antiaggregant; haemostatic; vulnerary;
CC antiulcer; osteopathic; dermatological; antiallergic; antiasthmatic;
CC antidiabetic; cytostatic; neuroprotective; antidepressant; nootropic;
CC antiparkinsonian; and immunostimulant. The proteins and polynucleotides
CC encoding them can be used in gene therapy, antisense therapy and vaccine
CC production, The proteins and polynucleotides are useful for screening for
CC agonists or antagonists of a protein and for the treatment and diagnosis
CC of disorders associated with the activity of a protein e.g. inflammation,
CC rheumatoid arthritis, septic shock, pancreatitis, cardiac dysfunction,
CC neuropathology, cardiac anaphylaxis, viral, bacterial, HIV and fungal
CC infections, autoimmunity, genetic diseases, haematopoietic disorders,
CC anaemia, platelet disorders, thrombocytopaenia, wounds, burns, ulcers,
CC osteoporosis, severe combined immunodeficiency, eczema, allergic
CC rhinitis, asthma, diabetes, cancer, multiple sclerosis, depression,
CC Alzheimer's disease, Parkinson's disease, neurodegenerative and
CC neurological disorders
XX
SQ Sequence 387 AA;

Query Match 100.0%; Score 1115; DB 4; Length 387;
Best Local Similarity 100.0%; Pred. No. 2e-104;
Matches 206; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 ATGRLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKL 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         35 ATGRLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKL 94

Qy         61 IEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEPSCGSEV 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         95 IEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEPSCGSEV 154

Qy        121 CVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGEETELTTPV 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        155 CVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGEETELTTPV 214

Qy        181 LPEETQEEDAKKTFKESREAALNLAY 206
              |||||||||||||||||||||||||
Db        215 LPEETQEEDAKKTFKESREAALNLAY 240
```

RESULT 7
ADA54574
ID ADA54574 standard; protein; 374 AA.
XX
AC ADA54574;
XX
DT 20-NOV-2003 (first entry)
XX
DE Human protein, SEQ ID 2142.
XX
KW Cytostatic; Anti-inflammatory; Osteopathic; Neuroprotective; Nootropic;
KW Gene Therapy; human; secretory protein; membrane proteins; cancer;
KW inflammatory disease; osteoporosis; neurological disease.
XX
OS Homo sapiens.
XX
PN EP1293569-A2.
XX
PD 19-MAR-2003.
XX
PF 21-MAR-2002; 2002EP-00006586.
XX
PR 14-SEP-2001; 2001JP-00328381.
PR 24-JAN-2002; 2002US-0350435P.
XX
PA (HELI-) HELIX RES INST.
PA (REAS-) RES ASSOC BIOTECHNOLOGY.
XX
PI Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;
PI Yamamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Irie R, Tamechika I;
PI Seki N, Yoshikawa T, Otsuka M, Nagahari K, Masuho Y;
XX
DR WPI; 2003-395539/38.
DR N-PSDB; ADA52935.
XX
PT New polynucleotides encoding full-length polypeptides, e.g. secretory
PT and/or membrane proteins, useful for developing medicines for diseases in
PT which the gene is involved, or as target molecules for gene therapy.
XX
PS Claim 14; SEQ ID NO 2142; 205pp; English.
XX
CC The present invention relates to novel human secretory or membrane
CC proteins (ADA54072-ADA55710) and their coding sequences (ADA52433-
CC ADA54071). The coding sequences are useful in the gene therapy of
CC diseases caused by abnormalities of the proteins, e.g. cancer,
CC inflammatory diseases, osteoporosis or neurological disease.
XX
SQ Sequence 374 AA;

Query Match 99.6%; Score 1111; DB 6; Length 374;
Best Local Similarity 99.5%; Pred. No. 4.8e-104;
Matches 205; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 ATGRLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKL 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ATGRLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKL 81

Qy         61 IEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEPSCGSEV 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 IEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEPSCGSEV 141

Qy        121 CVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGEETELTTPV 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 CVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGEETELTTPV 201

Qy        181 LPEETQEEDAKKTFKESREAALNLAY 206
              ||||||||| ||||||||||||||||
Db        202 LPEETQEEDTKKTFKESREAALNLAY 227
```

RESULT 8
AAY91490
ID AAY91490 standard; protein; 374 AA.
XX
AC AAY91490;
XX
DT 29-JUN-2000 (first entry)
XX
DE Human secreted protein sequence encoded by gene 40 SEQ ID NO:163.
XX
KW Human; secreted protein; diagnosis; cytostatic; immunosuppressive;
KW antiHIV; antiinflammatory; nootropic; neuroprotective; antiallergic;
KW osteopathic; antiarthritic; antibacterial; antidiabetic; antiasthma;
KW antipsoriatic; cardiant; gene therapy; cancer; neurological disorder;
KW immune disease; inflammation; blood disorder; tumour.
XX
OS Homo sapiens.
XX
PN WO200006698-A1.
XX
PD 10-FEB-2000.
XX
PF 29-JUL-1999; 99WO-US017130.
XX
PR 30-JUL-1998; 98US-0094657P.

PR 05-AUG-1998; 98US-0095486P.
PR 06-AUG-1998; 98US-0095454P.
PR 06-AUG-1998; 98US-0095455P.
PR 12-AUG-1998; 98US-0096319P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Komatsoulis GA, Rosen CA, Ruben SM, Duan R, Moore PA, Shi Y;
PI Lafleur D, Wei Y, Ni J, Florence KA, Young PE, Brewer LA;
PI Soppet DR, Endress GA, Ebner R, Olsen HS, Mucenski M;
XX
DR WPI; 2000-195282/17.
DR N-PSDB; AAA26385.
XX
PT New isolated human genes and the secreted polypeptides they encode,
PT useful for diagnosis and treatment of e.g. cancers, neurological
PT disorders, immune diseases, inflammation or blood disorders.
XX
PS Claim 11; Page 483-484; 634pp; English.
XX
CC The polynucleotide sequences given in AAA26346 to AAA26458 encode the
CC human secreted proteins given in AAY91451 to AAY91691. The human secreted
CC proteins can have activities based on the tissues and cells they are
CC expressed in. Examples of the activities are: cytostatic;
CC immunosuppressive; antiHIV; antiinflammatory; nootropic; neuroprotective;
CC antiallergic; osteopathic; antiarthritic; antibacterial; antidiabetic;
CC antiasthma; antipsoriatic; and cardiant. The polynucleotides and their
CC corresponding secreted proteins are useful for preventing, treating or
CC ameliorating medical conditions, e.g. by protein or gene therapy. Also
CC pathological conditions can be diagnosed by determining the amount of the
CC proteins in a sample or by determining the presence of mutations in the
CC polynucleotides. Specific uses are described for each of the
CC polynucleotides, based on which tissues they are most highly expressed
CC in, and include developing products for the diagnosis or treatment of
CC cancer, tumours, neurodegenerative disorders, developmental abnormalities
CC and foetal deficiencies, blood disorders, diseases of the immune system,
CC autoimmune diseases, hepatic and renal disease, inflammation, allergies,
CC Alzheimer's and behavioural disorders, schizophrenia, osteoporosis,
CC arthritis, infections, AIDS, spinal cord injuries, transplant rejection,
CC diabetes, asthma, sepsis, acne, psoriasis, cardiovascular disorders,
CC reproductive disorders, gastrointestinal disorders, respiratory disorders
CC and metabolic disorders. The proteins or polynucleotides can also be used
CC as food additives or preservatives. The proteins are also useful for
CC identifying their binding partners. AAA26337 to AAA26345 and AAY91450 are
CC sequences used in the exemplification of the present invention
XX
SQ Sequence 374 AA;

Query Match 99.1%; Score 1105; DB 3; Length 374;
Best Local Similarity 99.0%; Pred. No. 2e-103;
Matches 204; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy          1 ATGRLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKL 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ATGRLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKL 81

Qy         61 IEXFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEPSCGSEV 120
              || |||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db         82 IEXFIENLLPSDGDFWIGLRRREEKQSNSTXCQDLYAWTDGSISQFRNWYVDEPSCGSEV 141

Qy        121 CVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGEETELTTPV 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 CVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGEETELTTPV 201

Qy        181 LPEETQEEDAKKTFKESREAALNLAY 206
              |||||||||||||||||||||||||
Db        202 LPEETQEEDAKKTFKESREAALNLAY 227
```

RESULT 9
AAY13367
ID AAY13367 standard; protein; 382 AA.
XX
AC AAY13367;
XX
DT 25-JUN-1999 (first entry)
XX
DE Amino acid sequence of protein PRO234.
XX
KW Secreted protein; transmembrane protein; human; enterocolitis;
KW Zollinger-Ellison syndrome; gastrointestinal ulceration;
KW congenital microvillus atrophy; skin disease; cell growth;
KW abnormal keratinocyte differentiation; psoriasis; epithelial cancer;
KW Parkinson's disease; Alzheimer's disease; ALS; neuropathy; fibromodulin;
KW dermal scarring; Usher Syndrome; Atrophia areata; anti-thrombotic;
KW wound healing; tissue repair.
XX
OS Homo sapiens.
XX
PN WO9914328-A2.
XX
PD 25-MAR-1999.
XX
PF 16-SEP-1998; 98WO-US019330.
XX
PR 17-SEP-1997; 97US-0059113P.
PR 17-SEP-1997; 97US-0059115P.
PR 17-SEP-1997; 97US-0059117P.
PR 17-SEP-1997; 97US-0059119P.
PR 17-SEP-1997; 97US-0059121P.
PR 17-SEP-1997; 97US-0059122P.
PR 17-SEP-1997; 97US-0059184P.
PR 18-SEP-1997; 97US-0059263P.
PR 18-SEP-1997; 97US-0059266P.
PR 15-OCT-1997; 97US-0062125P.
PR 17-OCT-1997; 97US-0062285P.
PR 17-OCT-1997; 97US-0062287P.
PR 21-OCT-1997; 97US-0063486P.
PR 24-OCT-1997; 97US-0062814P.
PR 24-OCT-1997; 97US-0062816P.
PR 24-OCT-1997; 97US-0063045P.
PR 24-OCT-1997; 97US-0063120P.
PR 24-OCT-1997; 97US-0063121P.
PR 24-OCT-1997; 97US-0063127P.
PR 24-OCT-1997; 97US-0063128P.
PR 27-OCT-1997; 97US-0063327P.
PR 27-OCT-1997; 97US-0063329P.
PR 28-OCT-1997; 97US-0063541P.
PR 28-OCT-1997; 97US-0063542P.
PR 28-OCT-1997; 97US-0063544P.
PR 28-OCT-1997; 97US-0063549P.
PR 28-OCT-1997; 97US-0063550P.
PR 28-OCT-1997; 97US-0063564P.
PR 29-OCT-1997; 97US-0063435P.
PR 29-OCT-1997; 97US-0063704P.
PR 29-OCT-1997; 97US-0063732P.
PR 29-OCT-1997; 97US-0063734P.
PR 29-OCT-1997; 97US-0063735P.
PR 29-OCT-1997; 97US-0063738P.
PR 29-OCT-1997; 97US-0064215P.
PR 31-OCT-1997; 97US-0063870P.
PR 31-OCT-1997; 97US-0064103P.
PR 03-NOV-1997; 97US-0064248P.
PR 07-NOV-1997; 97US-0064809P.
PR 12-NOV-1997; 97US-0065186P.
PR 17-NOV-1997; 97US-0065846P.
PR 18-NOV-1997; 97US-0065693P.
PR 21-NOV-1997; 97US-0066120P.
PR 21-NOV-1997; 97US-0066364P.
PR 24-NOV-1997; 97US-0066453P.
PR 24-NOV-1997; 97US-0066466P.
PR 24-NOV-1997; 97US-0066511P.
PR 24-NOV-1997; 97US-0066770P.
PR 24-NOV-1997; 97US-0066772P.
PR 25-NOV-1997; 97US-0066840P.
XX

PA (GETH) GENENTECH INC.
XX
PI Wood WI, Gurney AL, Goddard A, Pennica D, Chen J, Yuan J;
XX
DR WPI; 1999-229533/19.
DR N-PSDB; AAX52238.
XX
PT New isolated human genes and polypeptides used in, e.g. treatment of
PT gastrointestinal ulceration.
XX
PS Claim 12; Fig 50; 320pp; English.
XX
CC AAY13344-403 represent secreted and transmembrane human proteins. The
CC cDNA sequences are obtained from cDNA libraries, prepared from fetal
CC lung, fetal kidney, fetal brain, fetal liver and fetal retina. The
CC encoded polypeptides have specific uses based on their homology to known
CC polypeptides, e.g. PRO211 and PRO217 can be used for disorders associated
CC with the preservation and maintenance of gastrointestinal mucosa and the
CC repair of acute and chronic mucosal lesions (e.g. enterocolitis,
CC Zollinger-Ellison syndrome, gastrointestinal ulceration and congenital
CC microvillus atrophy), skin diseases associated with abnormal keratinocyte
CC differentiation (e.g. psoriasis, epithelial cancers such as lung squamous
CC cell carcinoma of the vulva and gliomas), potent effects on cell growth
CC and development, diseases related to growth or survival of nerve cells
CC including Parkinson's disease, Alzheimer's disease, ALS, neuropathies or
CC cancer. PRO265 can be used as for fibromodulin, e.g. for reducing dermal
CC scarring. PRO264 can be used as a target for anti-tumor drugs. PRO533 may
CC be used in the treatment of Usher Syndrome or Atrophia areata; PRO269 can
CC be used as an anti-thrombotic agent; PRO287 polypeptides and portions may
CC have therapeutic applications in wound healing and tissue repair; PRO317
CC can be used for treating problems of the kidney, uterus, endometrium,
CC blood vessels, or related tissue, e.g. in the heart of genital tract
XX
SQ Sequence 382 AA;

Query Match 98.7%; Score 1101; DB 2; Length 382;
Best Local Similarity 96.3%; Pred. No. 5.2e-103;
Matches 206; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

```
Qy          1 ATGRLLS--------GQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 52
              |||||||        ||||||||||||||||||||||||||||||||||||||||||||
Db         22 ATGRLLSASDLDLRGGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 81

Qy         53 ESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVD 112
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 ESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVD 141

Qy        113 EPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGE 172
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 EPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGE 201

Qy        173 ETELTTPVLPEETQEEDAKKTFKESREAALNLAY 206
              |||||||||||||||||||||||||||||||||
Db        202 ETELTTPVLPEETQEEDAKKTFKESREAALNLAY 235
```

RESULT 10
ADC78457
ID ADC78457 standard; protein; 382 AA.
XX
AC ADC78457;
XX
DT 01-JAN-2004 (first entry)
XX
DE Human PRO234 protein.
XX
KW antiinflamatory; antiulcer; cytostatic; antipsoriatic; antiparkinsonian;
KW nootropic; neuroprotective; vasotropic; chemotaxic; angiogenic;
KW neurotrophic; osteopathic; antiasthmatic; antiarthritic; antirheumatic;
KW antiarteriosclerotic; cardiant; antidiabetic; cerebroprotective;
KW thrombolytic; immunomodulator; enterocolitis; Zollinger-Ellison syndrome;
KW gastrointestinal ulceration; psoriasis; cancer; Parkinson's disease;
KW Alzheimer's; ALS; neuropathy; dermal scarring; wound healing;
KW nerve repair; thrombosis; bone; cartilage formation; angiogenesis;
KW asthma; rheumatoid arthritis; multiple sclerosis; inflammatory disorder;
KW atherosclerosis; cardiac injury; infertility; premature aging; AIDS;
KW diabetes; stroke; gene therapy; transgenic; PRO; human.
XX
OS Homo sapiens.
XX
PN WO200015796-A2.
XX
PD 23-MAR-2000.
XX
PF 15-SEP-1999; 99WO-US021090.
XX
PR 16-SEP-1998; 98WO-US019330.
XX
PA (GETH) GENENTECH INC.
XX
PI Chen J, Goddard A, Gurney AL, Hillan K, Pennica D, Wood WI;
PI Yuan J;
XX
DR WPI; 2000-271434/23.
DR N-PSDB; ADC78456.
XX
PT Novel nucleic acids encoding secreted and transmembrane polypeptides with
PT homology, e.g. to growth and cancer-associated antigens.
XX
PS Claim 12; SEQ ID NO 137; 355pp; English.
XX
CC The invention relates to a novel nucleic acid encoding a PRO polypeptide.
CC The polypeptides and polynucleotides of the invention may be useful as
CC research tools and as therapeutics for treating enterocolitis, Zollinger-
CC Ellison syndrome, gastrointestinal ulceration, psoriasis, cancer,
CC Parkinson's disease, Alzheimer's disease, ALS, neuropathies, dermal
CC scarring and wound healing, nerve repair, thrombosis, bone and/or
CC cartilage formation, angiogenesis, asthma, rheumatoid arthritis, multiple
CC sclerosis, inflammatory disorders, atherosclerosis, cardiac injury,
CC infertility, premature aging, AIDS, diabetes complications and stroke.
CC The molecules may also be utilised during gene therapy procedures and
CC transgenic animal production. The current sequence is that of the human
CC PRO protein of the invention.
XX
SQ Sequence 382 AA;

Query Match 98.7%; Score 1101; DB 3; Length 382;
Best Local Similarity 96.3%; Pred. No. 5.2e-103;
Matches 206; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

```
Qy          1 ATGRLLS--------GQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 52
              |||||||        ||||||||||||||||||||||||||||||||||||||||||||
Db         22 ATGRLLSASDLDLRGGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 81

Qy         53 ESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVD 112
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 ESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVD 141

Qy        113 EPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGE 172
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 EPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGE 201

Qy        173 ETELTTPVLPEETQEEDAKKTFKESREAALNLAY 206
              |||||||||||||||||||||||||||||||||
Db        202 ETELTTPVLPEETQEEDAKKTFKESREAALNLAY 235
```

RESULT 11
AAB80235
ID AAB80235 standard; protein; 382 AA.
XX
AC AAB80235;
XX
DT 24-APR-2001 (first entry)
XX
DE Human PRO234 protein.

```
XX
KW   Human; PRO; dermatological; antipsoriatic; cytostatic; antiinflammatory;
KW   antiparkinsonian nootropic; neuroprotective; vulnerary; cardiant;
KW   antiangiogenic; vasotropic; antiasthmatic; antirheumatic; cancer;
KW   antiarthritic; antiinfertility; antidiabetic; antiviral; diabetes;
KW   ophthalmological; gene therapy; skin disease; gastrointestinal disorder;
KW   ischaemia; inflammation.
XX
OS   Homo sapiens.
XX
PN   WO200104311-A1.
XX
PD   18-JAN-2001.
XX
PF   22-FEB-2000; 2000WO-US004414.
XX
PR   07-JUL-1999;   99US-0143048P.
PR   26-JUL-1999;   99US-0145698P.
PR   28-JUL-1999;   99US-0146222P.
PR   08-SEP-1999;   99WO-US020594.
PR   13-SEP-1999;   99WO-US020944.
PR   15-SEP-1999;   99WO-US021090.
PR   15-SEP-1999;   99WO-US021547.
PR   05-OCT-1999;   99WO-US023089.
PR   29-NOV-1999;   99WO-US028214.
PR   30-NOV-1999;   99WO-US028313.
PR   02-DEC-1999;   99WO-US028564.
PR   02-DEC-1999;   99WO-US028565.
PR   16-DEC-1999;   99WO-US030095.
PR   20-DEC-1999;   99WO-US030911.
PR   20-DEC-1999;   99WO-US030999.
PR   05-JAN-2000; 2000WO-US000219.
XX
PA   (GETH ) GENENTECH INC.
XX
PI   Ashkenazi AJ,  Botstein D,  Desnoyers L,  Eaton DL,  Ferrara N;
PI   Filvaroff E,  Fong S,  Gao W,  Gerber H,  Gerritsen ME,  Goddard A;
PI   Godowski PJ,  Grimaldi CJ,  Gurney AL,  Hillan KJ,  Kljavin IJ;
PI   Mather JP,  Pan J,  Paoni NF,  Roy MA,  Stewart TA,  Tumas D;
PI   Williams PM,  Wood WI;
XX
DR   WPI; 2001-081051/09.
DR   N-PSDB; AAF72396.
XX
PT   Sixty one nucleic acids encoding PRO polypeptides which are useful in the
PT   treatment of skin diseases (e.g. psoriasis), cancers (e.g. lung squamous
PT   cell carcinoma) and neurodegenerative diseases (e.g. Alzheimer's
PT   disease).
XX
PS   Claim 1; Fig 50; 393pp; English.
XX
CC   The present sequence is one of sixty one novel secreted and transmembrane
CC   PRO polypeptides. The PRO polypeptides are useful for treating skin
CC   diseases (e.g. psoriasis), cancers (e.g. lung squamous cell carcinoma),
CC   gastrointestinal disorders (e.g. enterocolitis), neurodegenerative
CC   diseases (e.g. Alzheimer's disease, Parkinson's disease), wound repair,
CC   cardiovascular disorders (e.g. endometrial bleeding angiogenesis,
CC   ischaemias such as coronary ischaemia, atherosclerosis), inflammatory
CC   disorders (e.g. asthma, rheumatoid arthritis, multiple sclerosis),
CC   infertility, AIDS and diabetes and retinal disorders such as retinitis
CC   pigmentosum. The PRO nucleic acids have applications in molecular
CC   biology, including use as hybridization probes, and in chromosome and
CC   gene mapping
XX
SQ   Sequence 382 AA;

  Query Match             98.7%;  Score 1101;  DB 4;  Length 382;
  Best Local Similarity   96.3%;  Pred. No. 5.2e-103;
  Matches  206;  Conservative    0;  Mismatches    0;  Indels    8;  Gaps    1;

Qy          1 ATGRLLS--------GQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 52
              |||||||        ||||||||||||||||||||||||||||||||||||||||||||
Db         22 ATGRLLSASDLDLRGGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 81

Qy         53 ESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVD 112
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 ESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVD 141

Qy        113 EPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGE 172
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 EPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGE 201

Qy        173 ETELTTPVLPEETQEEDAKKTFKESREAALNLAY 206
              |||||||||||||||||||||||||||||||||
Db        202 ETELTTPVLPEETQEEDAKKTFKESREAALNLAY 235


RESULT 12
AAU29033
ID   AAU29033 standard; protein; 382 AA.
XX
AC   AAU29033;
XX
DT   18-DEC-2001  (first entry)
XX
DE   Human PRO polypeptide sequence #10.
XX
KW   PRO polypeptide; mammal; tumour; cancer; human; cattle; horse; sheep;
KW   dog; cat; pig; goat; rabbit; tumour necrosis factor alpha; TNF-alpha;
KW   blood; chondrocyte cell; cell proliferation; cell differentiation; colon;
KW   adrenal; lung; breast; prostate; rectum; cervix; liver; genetic disorder.
XX
OS   Homo sapiens.
XX
PN   WO200168848-A2.
XX
PD   20-SEP-2001.
XX
PF   28-FEB-2001; 2001WO-US006520.
XX
PR   01-MAR-2000; 2000WO-US005601.
PR   02-MAR-2000; 2000WO-US005841.
PR   03-MAR-2000; 2000US-0187202P.
PR   06-MAR-2000; 2000US-0186968P.
PR   14-MAR-2000; 2000US-0189320P.
PR   14-MAR-2000; 2000US-0189328P.
PR   15-MAR-2000; 2000WO-US006884.
PR   21-MAR-2000; 2000US-0190828P.
PR   21-MAR-2000; 2000US-0191007P.
PR   21-MAR-2000; 2000US-0191048P.
PR   21-MAR-2000; 2000US-0191314P.
PR   28-MAR-2000; 2000US-0192655P.
PR   29-MAR-2000; 2000US-0193032P.
PR   29-MAR-2000; 2000US-0193053P.
PR   30-MAR-2000; 2000WO-US008439.
PR   04-APR-2000; 2000US-0194449P.
PR   04-APR-2000; 2000US-0194647P.
PR   11-APR-2000; 2000US-0195975P.
PR   11-APR-2000; 2000US-0196000P.
PR   11-APR-2000; 2000US-0196187P.
PR   11-APR-2000; 2000US-0196690P.
PR   11-APR-2000; 2000US-0196820P.
PR   18-APR-2000; 2000US-0198121P.
PR   18-APR-2000; 2000US-0198585P.
PR   25-APR-2000; 2000US-0199397P.
PR   25-APR-2000; 2000US-0199550P.
PR   25-APR-2000; 2000US-0199654P.
PR   03-MAY-2000; 2000US-0201516P.
PR   17-MAY-2000; 2000WO-US013705.
PR   22-MAY-2000; 2000WO-US014042.
PR   30-MAY-2000; 2000WO-US014941.
PR   02-JUN-2000; 2000WO-US015264.
PR   05-JUN-2000; 2000US-0209832P.
PR   28-JUL-2000; 2000WO-US020710.
PR   22-AUG-2000; 2000US-00644848.
PR   24-AUG-2000; 2000WO-US023328.
```

PR 08-NOV-2000; 2000WO-US030952.
PR 01-DEC-2000; 2000WO-US032678.
PR 20-DEC-2000; 2000WO-US034956.
XX
PA (GETH) GENENTECH INC.
XX
PI Baker KP, Chen J, Desnoyers L, Goddard A, Godowski PJ, Gurney AL;
PI Pan J, Smith V, Watanabe CK, Wood WI, Zhang Z;
XX
DR WPI; 2001-602746/68.
DR N-PSDB; AAS45934.
XX
PT Novel nucleic acids encoding PRO polypeptides, used to diagnose the
PT presence of tumors, such as prostate and breast tumors, in mammals and to
PT screen for modulators of the compounds.
XX
PS Claim 11; Fig 20; 774pp; English.
XX
CC Sequences AAU29024-AAU29328 represent PRO polypeptides of the invention.
CC The PRO polypeptides and their associated nucleic acids can be used to
CC detect the presence of a tumour in a mammal by comparing the level of
CC expression of a PRO polypeptide in a test sample of cells from the animal
CC and a control sample of normal cells, whereby a higher level of
CC expression in the test sample indicates the presence of a tumour in the
CC mammal. Mammals include dogs, cats, cattle, horses, sheep, pigs, goats
CC and rabbits but are preferably human. The polypeptides can be used to
CC stimulate tumour necrosis factor (TNF) alpha release from human blood,
CC when contacted with it. A specific polypeptide can be used to stimulate
CC the proliferation or differentiation of chondrocyte cells. The PRO
CC proteins can be used to determine the presence of tumours and also
CC susceptibility to tumour development, particularly adrenal, lung, colon,
CC breast, prostate, rectal, cervical, or liver tumours, in mammalian
CC subjects. The oligonucleotide probes specific for the PRO nucleic acids
CC can be used for genetic analysis of individuals with genetic disorders
XX
SQ Sequence 382 AA;

```
  Query Match             98.7%;  Score 1101;  DB 4;  Length 382;
  Best Local Similarity   96.3%;  Pred. No. 5.2e-103;
  Matches  206;  Conservative    0;  Mismatches    0;  Indels    8;  Gaps    1;

Qy          1 ATGRLLS--------GQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI  52
              |||||||        ||||||||||||||||||||||||||||||||||||||||||||
Db         22 ATGRLLSASDLDLRGGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI  81

Qy         53 ESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVD  112
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 ESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVD  141

Qy        113 EPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGE  172
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 EPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGE  201

Qy        173 ETELTTPVLPEETQEEDAKKTFKESREAALNLAY 206
              |||||||||||||||||||||||||||||||||
Db        202 ETELTTPVLPEETQEEDAKKTFKESREAALNLAY 235
```

RESULT 13
ABU58409
ID ABU58409 standard; protein; 382 AA.
XX
AC ABU58409;
XX
DT 15-APR-2003 (first entry)
XX
DE Human PRO polypeptide #10.
XX
KW Human; PRO; cytostatic; tumour; cancer; breast; lung; stomach; liver;
KW dog; cat; cow; horse; sheep; pig; goat; rabbit; ADEPT;
KW antibody-dependent enzyme mediated prodrug therapy.
XX
OS Homo sapiens.

XX
PN US2003027272-A1.
XX
PD 06-FEB-2003.
XX
PF 21-JUN-2002; 2002US-00176492.
XX
PR 18-SEP-1997; 97US-0059263P.
PR 18-SEP-1997; 97US-0059266P.
PR 17-OCT-1997; 97US-0062250P.
PR 21-OCT-1997; 97US-0063486P.
PR 24-OCT-1997; 97US-0063120P.
PR 24-OCT-1997; 97US-0063121P.
PR 28-OCT-1997; 97US-0063540P.
PR 28-OCT-1997; 97US-0063541P.
PR 28-OCT-1997; 97US-0063544P.
PR 28-OCT-1997; 97US-0063564P.
PR 29-OCT-1997; 97US-0063734P.
PR 31-OCT-1997; 97US-0063870P.
PR 31-OCT-1997; 97US-0064103P.
PR 13-NOV-1997; 97US-0065311P.
PR 21-NOV-1997; 97US-0066120P.
PR 24-NOV-1997; 97US-0066466P.
PR 24-NOV-1997; 97US-0066772P.
PR 11-DEC-1997; 97US-0069335P.
PR 12-DEC-1997; 97US-0069425P.
PR 17-DEC-1997; 97US-0069870P.
PR 18-DEC-1997; 97US-0068017P.
PR 10-MAR-1998; 98US-0077450P.
PR 11-MAR-1998; 98US-0077632P.
PR 11-MAR-1998; 98US-0077649P.
PR 20-MAR-1998; 98US-0078886P.
PR 20-MAR-1998; 98US-0078939P.
PR 27-MAR-1998; 98US-0079664P.
PR 27-MAR-1998; 98US-0079786P.
PR 31-MAR-1998; 98US-0080107P.
PR 31-MAR-1998; 98US-0080194P.
PR 01-APR-1998; 98US-0080327P.
PR 01-APR-1998; 98US-0080333P.
PR 08-APR-1998; 98US-0081049P.
PR 08-APR-1998; 98US-0081070P.
PR 09-APR-1998; 98US-0081195P.
PR 15-APR-1998; 98US-0081838P.
PR 21-APR-1998; 98US-0082568P.
PR 21-APR-1998; 98US-0082569P.
PR 22-APR-1998; 98US-0082704P.
PR 22-APR-1998; 98US-0082797P.
PR 28-APR-1998; 98US-0083322P.
PR 29-APR-1998; 98US-0083495P.
PR 29-APR-1998; 98US-0083496P.
PR 29-APR-1998; 98US-0083499P.
PR 29-APR-1998; 98US-0083559P.
PR 05-MAY-1998; 98US-0084366P.
PR 06-MAY-1998; 98US-0084414P.
PR 07-MAY-1998; 98US-0084639P.
PR 07-MAY-1998; 98US-0084640P.
PR 07-MAY-1998; 98US-0084643P.
PR 15-MAY-1998; 98US-0085579P.
PR 15-MAY-1998; 98US-0085580P.
PR 15-MAY-1998; 98US-0085582P.
PR 15-MAY-1998; 98US-0085700P.
PR 18-MAY-1998; 98US-0086023P.
PR 22-MAY-1998; 98US-0086392P.
PR 22-MAY-1998; 98US-0086486P.
PR 28-MAY-1998; 98US-0087098P.
PR 28-MAY-1998; 98US-0087208P.
PR 02-JUN-1998; 98US-0087609P.
PR 02-JUN-1998; 98US-0087759P.
PR 03-JUN-1998; 98US-0087827P.
PR 04-JUN-1998; 98US-0088025P.
PR 04-JUN-1998; 98US-0088028P.
PR 04-JUN-1998; 98US-0088029P.
PR 04-JUN-1998; 98US-0088033P.

PR 04-JUN-1998; 98US-0088326P.
PR 05-JUN-1998; 98US-0088167P.
PR 05-JUN-1998; 98US-0088202P.
PR 05-JUN-1998; 98US-0088212P.
PR 05-JUN-1998; 98US-0088217P.
PR 09-JUN-1998; 98US-0088655P.
PR 10-JUN-1998; 98US-0088722P.
PR 10-JUN-1998; 98US-0088738P.
PR 10-JUN-1998; 98US-0088740P.
PR 10-JUN-1998; 98US-0088811P.
PR 10-JUN-1998; 98US-0088824P.
PR 10-JUN-1998; 98US-0088825P.
PR 10-JUN-1998; 98US-0088826P.
PR 11-JUN-1998; 98US-0088861P.
PR 11-JUN-1998; 98US-0088863P.
PR 11-JUN-1998; 98US-0088876P.
PR 12-JUN-1998; 98US-0089090P.
PR 12-JUN-1998; 98US-0089105P.
PR 16-JUN-1998; 98US-0089512P.
PR 16-JUN-1998; 98US-0089514P.
PR 17-JUN-1998; 98US-0089538P.
PR 17-JUN-1998; 98US-0089598P.
PR 17-JUN-1998; 98US-0089653P.
PR 18-JUN-1998; 98US-0089908P.
PR 19-JUN-1998; 98US-0089952P.
PR 22-JUN-1998; 98US-0090246P.
PR 22-JUN-1998; 98US-0090252P.
PR 22-JUN-1998; 98US-0090254P.
PR 24-JUN-1998; 98US-0090429P.
PR 24-JUN-1998; 98US-0090435P.
PR 24-JUN-1998; 98US-0090444P.
PR 24-JUN-1998; 98US-0090461P.
PR 24-JUN-1998; 98US-0090535P.
PR 24-JUN-1998; 98US-0090540P.
PR 25-JUN-1998; 98US-0090676P.
PR 25-JUN-1998; 98US-0090678P.
PR 25-JUN-1998; 98US-0090688P.
PR 25-JUN-1998; 98US-0090690P.
PR 25-JUN-1998; 98US-0090694P.
PR 25-JUN-1998; 98US-0090695P.
PR 25-JUN-1998; 98US-0090696P.
PR 26-JUN-1998; 98US-00105413.
PR 26-JUN-1998; 98US-0090862P.
PR 26-JUN-1998; 98US-0090863P.
PR 26-JUN-1998; 98US-0091010P.
PR 01-JUL-1998; 98US-0091359P.
PR 01-JUL-1998; 98US-0091544P.
PR 02-JUL-1998; 98US-0091478P.
PR 02-JUL-1998; 98US-0091486P.
PR 02-JUL-1998; 98US-0091626P.
PR 02-JUL-1998; 98US-0091628P.
PR 02-JUL-1998; 98US-0091632P.
PR 24-JUL-1998; 98US-0094006P.
PR 04-AUG-1998; 98US-0095282P.
PR 10-AUG-1998; 98US-0095998P.
PR 10-AUG-1998; 98US-0096012P.
PR 17-AUG-1998; 98US-0096757P.
PR 17-AUG-1998; 98US-0096766P.
PR 17-AUG-1998; 98US-0096867P.
PR 17-AUG-1998; 98US-0096891P.
PR 17-AUG-1998; 98US-0096897P.
PR 18-AUG-1998; 98US-0096949P.
PR 18-AUG-1998; 98US-0096959P.
PR 18-AUG-1998; 98US-0097022P.
PR 26-AUG-1998; 98US-0097952P.
PR 26-AUG-1998; 98US-0097954P.
PR 26-AUG-1998; 98US-0097955P.
PR 26-AUG-1998; 98US-0097971P.
PR 26-AUG-1998; 98US-0097974P.
PR 26-AUG-1998; 98US-0098014P.
PR 01-SEP-1998; 98US-0098716P.
PR 01-SEP-1998; 98US-0098723P.
PR 02-SEP-1998; 98US-0098803P.
PR 02-SEP-1998; 98US-0098821P.
PR 02-SEP-1998; 98US-0098843P.
PR 09-SEP-1998; 98US-0099602P.
PR 10-SEP-1998; 98US-0099741P.
PR 10-SEP-1998; 98US-0099754P.
PR 10-SEP-1998; 98US-0099763P.
PR 10-SEP-1998; 98US-0099812P.
PR 15-SEP-1998; 98US-0100388P.
PR 16-SEP-1998; 98US-0100662P.
PR 16-SEP-1998; 98US-0100664P.
PR 16-SEP-1998; 98US-0101751P.
PR 16-SEP-1998; 98WO-US019330.
PR 17-SEP-1998; 98US-0100683P.
PR 17-SEP-1998; 98US-0100684P.
PR 17-SEP-1998; 98US-0100919P.
PR 17-SEP-1998; 98US-0100930P.
PR 18-SEP-1998; 98US-0100849P.
PR 18-SEP-1998; 98US-0101014P.
PR 18-SEP-1998; 98US-0101068P.
PR 23-SEP-1998; 98US-0101471P.
PR 23-SEP-1998; 98US-0101472P.
PR 23-SEP-1998; 98US-0101475P.
PR 23-SEP-1998; 98US-0101477P.
PR 24-SEP-1998; 98US-0101738P.
PR 24-SEP-1998; 98US-0101739P.
PR 24-SEP-1998; 98US-0101743P.
PR 24-SEP-1998; 98US-0101922P.
PR 25-SEP-1998; 98US-0101786P.
PR 29-SEP-1998; 98US-0102207P.
PR 29-SEP-1998; 98US-0102240P.
PR 29-SEP-1998; 98US-0102330P.
PR 29-SEP-1998; 98US-0102331P.
PR 30-SEP-1998; 98US-0102487P.
PR 30-SEP-1998; 98US-0102570P.
PR 30-SEP-1998; 98US-0102571P.
PR 01-OCT-1998; 98US-0102684P.
PR 01-OCT-1998; 98US-0102687P.
PR 02-OCT-1998; 98US-0102965P.
PR 06-OCT-1998; 98US-0103258P.
PR 06-OCT-1998; 98US-0103449P.
PR 07-OCT-1998; 98US-00168978.

Query Match 98.7%; Score 1101; DB 6; Length 382;
Best Local Similarity 96.3%; Pred. No. 5.2e-103;
Matches 206; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

```
Qy          1 ATGRLLS--------GQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 52
              |||||||         ||||||||||||||||||||||||||||||||||||||||||||
Db         22 ATGRLLSASDLDLRGGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 81

Qy         53 ESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVD 112
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 ESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVD 141

Qy        113 EPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGE 172
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 EPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGE 201

Qy        173 ETELTTPVLPEETQEEDAKKTFKESREAALNLAY 206
              |||||||||||||||||||||||||||||||||
Db        202 ETELTTPVLPEETQEEDAKKTFKESREAALNLAY 235
```

RESULT 14
ABU71613
ID ABU71613 standard; protein; 382 AA.
XX
AC ABU71613;
XX
DT 16-JUN-2003 (first entry)
XX
DE Human PRO polypeptide #24.
XX

KW Human; PRO; secreted polypeptide; transmembrane polypeptide;
KW pathological disorder; cardiac insufficiency disorder; protein secretion;
KW pancreas; diabetes; gastrointestinal mucosa; mucosal lesion; psoriasis;
KW skin disease; keratinocyte differentiation; epithelial cancer; tumour;
KW lung squamous cell carcinoma; epidermoid carcinoma; vulva; glioma;
KW cytostatic; cardiant; endocrine; antidiabetic; gastrointestinal;
KW antiulcer; dermatological; vulnerary.
XX
OS Homo sapiens.
XX
PN US2002146709-A1.
XX
PD 10-OCT-2002.
XX
PF 18-JUL-2001; 2001US-00909088.
XX
PR 17-SEP-1997; 97US-0059113P.
PR 17-SEP-1997; 97US-0059115P.
PR 17-SEP-1997; 97US-0059117P.
PR 17-SEP-1997; 97US-0059119P.
PR 17-SEP-1997; 97US-0059121P.
PR 17-SEP-1997; 97US-0059122P.
PR 17-SEP-1997; 97US-0059184P.
PR 18-SEP-1997; 97US-0059263P.
PR 18-SEP-1997; 97US-0059266P.
PR 15-OCT-1997; 97US-0062125P.
PR 17-OCT-1997; 97US-0062285P.
PR 17-OCT-1997; 97US-0062287P.
PR 21-OCT-1997; 97US-0063486P.
PR 24-OCT-1997; 97US-0062814P.
PR 24-OCT-1997; 97US-0062816P.
PR 24-OCT-1997; 97US-0063045P.
PR 24-OCT-1997; 97US-0063120P.
PR 24-OCT-1997; 97US-0063121P.
PR 24-OCT-1997; 97US-0063127P.
PR 24-OCT-1997; 97US-0063128P.
PR 27-OCT-1997; 97US-0063327P.
PR 27-OCT-1997; 97US-0063329P.
PR 28-OCT-1997; 97US-0063541P.
PR 28-OCT-1997; 97US-0063542P.
PR 28-OCT-1997; 97US-0063544P.
PR 28-OCT-1997; 97US-0063549P.
PR 28-OCT-1997; 97US-0063550P.
PR 28-OCT-1997; 97US-0063564P.
PR 29-OCT-1997; 97US-0063435P.
PR 29-OCT-1997; 97US-0063704P.
PR 29-OCT-1997; 97US-0063732P.
PR 29-OCT-1997; 97US-0063734P.
PR 29-OCT-1997; 97US-0063735P.
PR 29-OCT-1997; 97US-0063738P.
PR 29-OCT-1997; 97US-0064215P.
PR 31-OCT-1997; 97US-0063870P.
PR 31-OCT-1997; 97US-0064103P.
PR 03-NOV-1997; 97US-0064248P.
PR 07-NOV-1997; 97US-0064809P.
PR 12-NOV-1997; 97US-0065186P.
PR 17-NOV-1997; 97US-0065846P.
PR 18-NOV-1997; 97US-0065693P.
PR 21-NOV-1997; 97US-0066120P.
PR 21-NOV-1997; 97US-0066364P.
PR 24-NOV-1997; 97US-0066453P.
PR 24-NOV-1997; 97US-0066466P.
PR 24-NOV-1997; 97US-0066511P.
PR 24-NOV-1997; 97US-0066770P.
PR 24-NOV-1997; 97US-0066772P.
PR 10-SEP-1998; 98WO-US018824.
PR 14-SEP-1998; 98WO-US019177.
PR 16-SEP-1998; 98WO-US019330.
PR 17-SEP-1998; 98WO-US019437.
PR 01-DEC-1998; 98WO-US025108.
PR 08-SEP-1999; 99WO-US020594.
PR 13-SEP-1999; 99WO-US020944.
PR 15-SEP-1999; 99WO-US021090.
PR 15-SEP-1999; 99WO-US021547.
PR 05-OCT-1999; 99WO-US023089.
PR 29-NOV-1999; 99WO-US028214.
PR 30-NOV-1999; 99WO-US028313.
PR 01-DEC-1999; 99WO-US028301.
PR 02-DEC-1999; 99WO-US028564.
PR 02-DEC-1999; 99WO-US028565.
PR 16-DEC-1999; 99WO-US030095.
PR 20-DEC-1999; 99WO-US030911.
PR 20-DEC-1999; 99WO-US030999.
PR 05-JAN-2000; 2000WO-US000219.
PR 11-FEB-2000; 2000WO-US003565.
PR 22-FEB-2000; 2000WO-US004414.
PR 24-FEB-2000; 2000WO-US005004.
PR 02-MAR-2000; 2000WO-US005841.
PR 20-MAR-2000; 2000WO-US007377.
PR 30-MAR-2000; 2000WO-US008439.
PR 22-MAY-2000; 2000WO-US014042.
PR 02-JUN-2000; 2000WO-US015264.
PR 28-JUL-2000; 2000WO-US020710.
PR 24-AUG-2000; 2000WO-US023328.
PR 18-SEP-2000; 2000US-00665350.
XX
PA (GETH) GENENTECH INC.
XX
PI Ashkenazi A, Botstein D, Desnoyers L, Eaton DL, Ferrara N;
PI Filvaroff E, Fong S, Gao W, Gerber H, Gerritsen ME, Goddard A;
PI Godowski PJ, Grimaldi JC, Gurney AL, Hillan KJ, Kljavin IJ;
PI Mather JP, Pan J, Paoni NF, Roy MA, Stewart TA, Tumas D;
PI Williams PM, Wood WI;
XX
DR WPI; 2003-328338/31.
DR N-PSDB; ACA59008.
XX
PT Isolated nucleic acid useful for e.g., treating pathological disorders
PT encodes a secreted or transmembrane protein.
XX
PS Claim 12; Fig 50; 473pp; English.
XX
CC The invention relates to human PRO polypeptides (secreted or
CC transmembrane polypeptides) and the polynucleotides encoding them. The
CC PRO polypeptides and polynucleotides can be used in treating pathological
CC disorders and tumours, in therapeutic treatment of cardiac insufficiency
CC disorders and in therapeutic treatment of disorders involving protein
CC secretion by the pancreas, including diabetes. They can also be used in
CC treating disorders associated with the preservation and maintenance of
CC gastrointestinal mucosa and the repair of acute and chronic mucosal
CC lesions, and skin diseases associated with abnormal keratinocyte
CC differentiation (e.g., psoriases, epithelial cancers such as lung
CC squamous cell carcinoma, epidermoid carcinoma of the vulva and gliomas).
CC The sequences can be used as molecular markers for protein
CC electrophoresis purposes and can be utilised in protein-protein binding
CC assays, biochemical screening assays, immunoassays and cell-based assays.
CC This sequence represents a human PRO polypeptide of the invention
XX
SQ Sequence 382 AA;

Query Match 98.7%; Score 1101; DB 6; Length 382;
Best Local Similarity 96.3%; Pred. No. 5.2e-103;
Matches 206; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

```
Qy          1 ATGRLLS--------GQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 52
              |||||||        ||||||||||||||||||||||||||||||||||||||||||||
Db         22 ATGRLLSASDLDLRGGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 81

Qy         53 ESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVD 112
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 ESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVD 141

Qy        113 EPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGE 172
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 EPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGE 201
```

```
Qy        173 ETELTTPVLPEETQEEDAKKTFKESREAALNLAY 206
              |||||||||||||||||||||||||||||||||||
Db        202 ETELTTPVLPEETQEEDAKKTFKESREAALNLAY 235


RESULT 15
ABU87957
ID   ABU87957 standard; protein; 382 AA.
XX
AC   ABU87957;
XX
DT   07-JUL-2003  (first entry)
XX
DE   Novel human secreted and transmembrane protein PRO234.
XX
KW   Human; secreted and transmembrane protein: PRO; gene therapy;
KW   tumour necrosis factor-alpha release; TNF-alpha release;
KW   chondrocyte proliferation; chondrocyte differentiation; tumour;
KW   adrenal tumour; lung tumour; colon tumour; breast tumour;
KW   prostate tumour; rectal tumour; cervical tumour; liver tumour.
XX
OS   Homo sapiens.
XX
PN   US2003032127-A1.
XX
PD   13-FEB-2003.
XX
PF   26-JUN-2002; 2002US-00183012.
XX
PR   18-SEP-1997;   97US-0059263P.
PR   18-SEP-1997;   97US-0059266P.
PR   17-OCT-1997;   97US-0062250P.
PR   21-OCT-1997;   97US-0063486P.
PR   24-OCT-1997;   97US-0063120P.
PR   24-OCT-1997;   97US-0063121P.
PR   28-OCT-1997;   97US-0063540P.
PR   28-OCT-1997;   97US-0063541P.
PR   28-OCT-1997;   97US-0063544P.
PR   28-OCT-1997;   97US-0063564P.
PR   29-OCT-1997;   97US-0063734P.
PR   31-OCT-1997;   97US-0063870P.
PR   31-OCT-1997;   97US-0064103P.
PR   13-NOV-1997;   97US-0065311P.
PR   21-NOV-1997;   97US-0066120P.
PR   24-NOV-1997;   97US-0066466P.
PR   24-NOV-1997;   97US-0066772P.
PR   11-DEC-1997;   97US-0069335P.
PR   12-DEC-1997;   97US-0069425P.
PR   17-DEC-1997;   97US-0069870P.
PR   18-DEC-1997;   97US-0068017P.
PR   10-MAR-1998;   98US-0077450P.
PR   11-MAR-1998;   98US-0077632P.
PR   11-MAR-1998;   98US-0077649P.
PR   20-MAR-1998;   98US-0078886P.
PR   20-MAR-1998;   98US-0078939P.
PR   27-MAR-1998;   98US-0079664P.
PR   27-MAR-1998;   98US-0079786P.
PR   31-MAR-1998;   98US-0080107P.
PR   31-MAR-1998;   98US-0080194P.
PR   01-APR-1998;   98US-0080327P.
PR   01-APR-1998;   98US-0080333P.
PR   08-APR-1998;   98US-0081049P.
PR   08-APR-1998;   98US-0081070P.
PR   09-APR-1998;   98US-0081195P.
PR   15-APR-1998;   98US-0081838P.
PR   21-APR-1998;   98US-0082568P.
PR   21-APR-1998;   98US-0082569P.
PR   22-APR-1998;   98US-0082704P.
PR   22-APR-1998;   98US-0082797P.
PR   28-APR-1998;   98US-0083322P.
PR   29-APR-1998;   98US-0083495P.
PR   29-APR-1998;   98US-0083496P.
PR   29-APR-1998;   98US-0083499P.
PR   29-APR-1998;   98US-0083559P.
PR   05-MAY-1998;   98US-0084366P.
PR   06-MAY-1998;   98US-0084414P.
PR   07-MAY-1998;   98US-0084639P.
PR   07-MAY-1998;   98US-0084640P.
PR   07-MAY-1998;   98US-0084643P.
PR   15-MAY-1998;   98US-0085579P.
PR   15-MAY-1998;   98US-0085580P.
PR   15-MAY-1998;   98US-0085582P.
PR   15-MAY-1998;   98US-0085700P.
PR   18-MAY-1998;   98US-0086023P.
PR   22-MAY-1998;   98US-0086392P.
PR   22-MAY-1998;   98US-0086486P.
PR   28-MAY-1998;   98US-0087098P.
PR   28-MAY-1998;   98US-0087208P.
PR   02-JUN-1998;   98US-0087609P.
PR   02-JUN-1998;   98US-0087759P.
PR   03-JUN-1998;   98US-0087827P.
PR   04-JUN-1998;   98US-0088025P.
PR   04-JUN-1998;   98US-0088028P.
PR   04-JUN-1998;   98US-0088029P.
PR   04-JUN-1998;   98US-0088033P.
PR   04-JUN-1998;   98US-0088326P.
PR   05-JUN-1998;   98US-0088167P.
PR   05-JUN-1998;   98US-0088202P.
PR   05-JUN-1998;   98US-0088212P.
PR   05-JUN-1998;   98US-0088217P.
PR   09-JUN-1998;   98US-0088655P.
PR   10-JUN-1998;   98US-0088722P.
PR   10-JUN-1998;   98US-0088738P.
PR   10-JUN-1998;   98US-0088740P.
PR   10-JUN-1998;   98US-0088811P.
PR   10-JUN-1998;   98US-0088824P.
PR   10-JUN-1998;   98US-0088825P.
PR   10-JUN-1998;   98US-0088826P.
PR   11-JUN-1998;   98US-0088861P.
PR   11-JUN-1998;   98US-0088863P.
PR   11-JUN-1998;   98US-0088876P.
PR   12-JUN-1998;   98US-0089090P.
PR   12-JUN-1998;   98US-0089105P.
PR   16-JUN-1998;   98US-0089512P.
PR   16-JUN-1998;   98US-0089514P.
PR   17-JUN-1998;   98US-0089538P.
PR   17-JUN-1998;   98US-0089598P.
PR   17-JUN-1998;   98US-0089653P.
PR   18-JUN-1998;   98US-0089908P.
PR   19-JUN-1998;   98US-0089952P.
PR   22-JUN-1998;   98US-0090246P.
PR   22-JUN-1998;   98US-0090252P.
PR   22-JUN-1998;   98US-0090254P.
PR   24-JUN-1998;   98US-0090429P.
PR   24-JUN-1998;   98US-0090435P.
PR   24-JUN-1998;   98US-0090444P.
PR   24-JUN-1998;   98US-0090461P.
PR   24-JUN-1998;   98US-0090535P.
PR   24-JUN-1998;   98US-0090540P.
PR   25-JUN-1998;   98US-0090676P.
PR   25-JUN-1998;   98US-0090678P.
PR   25-JUN-1998;   98US-0090688P.
PR   25-JUN-1998;   98US-0090690P.
PR   25-JUN-1998;   98US-0090694P.
PR   25-JUN-1998;   98US-0090695P.
PR   25-JUN-1998;   98US-0090696P.
PR   26-JUN-1998;   98US-00105413.
PR   26-JUN-1998;   98US-0090862P.
PR   26-JUN-1998;   98US-0090863P.
PR   26-JUN-1998;   98US-0091010P.
PR   01-JUL-1998;   98US-0091359P.
PR   01-JUL-1998;   98US-0091544P.
PR   02-JUL-1998;   98US-0091478P.
PR   02-JUL-1998;   98US-0091486P.
PR   02-JUL-1998;   98US-0091626P.
PR   02-JUL-1998;   98US-0091628P.
```

PR 02-JUL-1998; 98US-0091632P.
PR 24-JUL-1998; 98US-0094006P.
PR 04-AUG-1998; 98US-0095282P.
PR 10-AUG-1998; 98US-0095998P.
PR 10-AUG-1998; 98US-0096012P.
PR 17-AUG-1998; 98US-0096757P.
PR 17-AUG-1998; 98US-0096766P.
PR 17-AUG-1998; 98US-0096867P.
PR 17-AUG-1998; 98US-0096891P.
PR 17-AUG-1998; 98US-0096897P.
PR 18-AUG-1998; 98US-0096949P.
PR 18-AUG-1998; 98US-0096959P.
PR 18-AUG-1998; 98US-0097022P.
PR 26-AUG-1998; 98US-0097952P.
PR 26-AUG-1998; 98US-0097954P.
PR 26-AUG-1998; 98US-0097955P.
PR 26-AUG-1998; 98US-0097971P.
PR 26-AUG-1998; 98US-0097974P.
PR 26-AUG-1998; 98US-0098014P.
PR 01-SEP-1998; 98US-0098716P.
PR 01-SEP-1998; 98US-0098723P.
PR 02-SEP-1998; 98US-0098803P.
PR 02-SEP-1998; 98US-0098821P.
PR 02-SEP-1998; 98US-0098843P.
PR 09-SEP-1998; 98US-0099602P.
PR 10-SEP-1998; 98US-0099741P.
PR 10-SEP-1998; 98US-0099754P.
PR 10-SEP-1998; 98US-0099763P.
PR 10-SEP-1998; 98US-0099812P.
PR 15-SEP-1998; 98US-0100388P.
PR 16-SEP-1998; 98US-0100662P.
PR 16-SEP-1998; 98US-0100664P.
PR 16-SEP-1998; 98US-0101751P.
PR 16-SEP-1998; 98WO-US019330.
PR 17-SEP-1998; 98US-0100683P.
PR 17-SEP-1998; 98US-0100684P.
PR 17-SEP-1998; 98US-0100919P.
PR 17-SEP-1998; 98US-0100930P.
PR 18-SEP-1998; 98US-0100849P.
PR 18-SEP-1998; 98US-0101014P.
PR 18-SEP-1998; 98US-0101068P.
PR 23-SEP-1998; 98US-0101471P.
PR 23-SEP-1998; 98US-0101472P.
PR 23-SEP-1998; 98US-0101475P.
PR 23-SEP-1998; 98US-0101477P.
PR 24-SEP-1998; 98US-0101738P.
PR 24-SEP-1998; 98US-0101739P.
PR 24-SEP-1998; 98US-0101743P.
PR 24-SEP-1998; 98US-0101922P.
PR 25-SEP-1998; 98US-0101786P.
PR 29-SEP-1998; 98US-0102207P.
PR 29-SEP-1998; 98US-0102240P.
PR 29-SEP-1998; 98US-0102330P.
PR 29-SEP-1998; 98US-0102331P.
PR 30-SEP-1998; 98US-0102487P.
PR 30-SEP-1998; 98US-0102570P.
PR 30-SEP-1998; 98US-0102571P.
PR 01-OCT-1998; 98US-0102684P.
PR 01-OCT-1998; 98US-0102687P.
PR 02-OCT-1998; 98US-0102965P.
PR 06-OCT-1998; 98US-0103258P.

Query Match 98.7%; Score 1101; DB 6; Length 382;
Best Local Similarity 96.3%; Pred. No. 5.2e-103;
Matches 206; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

```
Qy          1 ATGRLLS--------GQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI  52
              |||||||        ||||||||||||||||||||||||||||||||||||||||||||
Db         22 ATGRLLSASDLDLRGGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI  81

Qy         53 ESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVD  112
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 ESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVD  141

Qy        113 EPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGE  172
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 EPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGE  201

Qy        173 ETELTTPVLPEETQEEDAKKTFKESREAALNLAY  206
              |||||||||||||||||||||||||||||||||
Db        202 ETELTTPVLPEETQEEDAKKTFKESREAALNLAY  235
```

RESULT 16
ABU84272
ID ABU84272 standard; protein; 382 AA.
XX
AC ABU84272;
XX
DT 02-AUG-2003 (first entry)
XX
DE Human secreted/transmembrane protein (PRO) #10.
XX
KW Human; secreted and transmembrane protein; PRO; TNF-alpha;
KW tumour necrosis factor alpha; chondrocyte cell; tumour; gene therapy;
KW tissue typing.
XX
OS Homo sapiens.
XX
PN US2003032112-A1.
XX
PD 13-FEB-2003.
XX
PF 21-JUN-2002; 2002US-00176756.
XX
PR 18-SEP-1997; 97US-0059263P.
PR 18-SEP-1997; 97US-0059266P.
PR 17-OCT-1997; 97US-0062250P.
PR 21-OCT-1997; 97US-0063486P.
PR 24-OCT-1997; 97US-0063120P.
PR 24-OCT-1997; 97US-0063121P.
PR 28-OCT-1997; 97US-0063540P.
PR 28-OCT-1997; 97US-0063541P.
PR 28-OCT-1997; 97US-0063544P.
PR 28-OCT-1997; 97US-0063564P.
PR 29-OCT-1997; 97US-0063734P.
PR 31-OCT-1997; 97US-0063870P.
PR 31-OCT-1997; 97US-0064103P.
PR 13-NOV-1997; 97US-0065311P.
PR 21-NOV-1997; 97US-0066120P.
PR 24-NOV-1997; 97US-0066466P.
PR 24-NOV-1997; 97US-0066772P.
PR 11-DEC-1997; 97US-0069335P.
PR 12-DEC-1997; 97US-0069425P.
PR 17-DEC-1997; 97US-0069870P.
PR 18-DEC-1997; 97US-0068017P.
PR 10-MAR-1998; 98US-0077450P.
PR 11-MAR-1998; 98US-0077632P.
PR 11-MAR-1998; 98US-0077649P.
PR 20-MAR-1998; 98US-0078886P.
PR 20-MAR-1998; 98US-0078939P.
PR 27-MAR-1998; 98US-0079664P.
PR 27-MAR-1998; 98US-0079786P.
PR 31-MAR-1998; 98US-0080107P.
PR 31-MAR-1998; 98US-0080194P.
PR 01-APR-1998; 98US-0080327P.
PR 01-APR-1998; 98US-0080333P.
PR 08-APR-1998; 98US-0081049P.
PR 08-APR-1998; 98US-0081070P.
PR 09-APR-1998; 98US-0081195P.
PR 15-APR-1998; 98US-0081838P.
PR 21-APR-1998; 98US-0082568P.
PR 21-APR-1998; 98US-0082569P.
PR 22-APR-1998; 98US-0082704P.
PR 22-APR-1998; 98US-0082797P.
PR 28-APR-1998; 98US-0083322P.

PR 29-APR-1998; 98US-0083495P.
PR 29-APR-1998; 98US-0083496P.
PR 29-APR-1998; 98US-0083499P.
PR 29-APR-1998; 98US-0083559P.
PR 05-MAY-1998; 98US-0084366P.
PR 06-MAY-1998; 98US-0084414P.
PR 07-MAY-1998; 98US-0084639P.
PR 07-MAY-1998; 98US-0084640P.
PR 07-MAY-1998; 98US-0084643P.
PR 15-MAY-1998; 98US-0085579P.
PR 15-MAY-1998; 98US-0085580P.
PR 15-MAY-1998; 98US-0085582P.
PR 15-MAY-1998; 98US-0085700P.
PR 18-MAY-1998; 98US-0086023P.
PR 22-MAY-1998; 98US-0086392P.
PR 22-MAY-1998; 98US-0086486P.
PR 28-MAY-1998; 98US-0087098P.
PR 28-MAY-1998; 98US-0087208P.
PR 02-JUN-1998; 98US-0087609P.
PR 02-JUN-1998; 98US-0087759P.
PR 03-JUN-1998; 98US-0087827P.
PR 04-JUN-1998; 98US-0088025P.
PR 04-JUN-1998; 98US-0088028P.
PR 04-JUN-1998; 98US-0088029P.
PR 04-JUN-1998; 98US-0088033P.
PR 04-JUN-1998; 98US-0088326P.
PR 05-JUN-1998; 98US-0088167P.
PR 05-JUN-1998; 98US-0088202P.
PR 05-JUN-1998; 98US-0088212P.
PR 05-JUN-1998; 98US-0088217P.
PR 09-JUN-1998; 98US-0088655P.
PR 10-JUN-1998; 98US-0088722P.
PR 10-JUN-1998; 98US-0088738P.
PR 10-JUN-1998; 98US-0088740P.
PR 10-JUN-1998; 98US-0088811P.
PR 10-JUN-1998; 98US-0088824P.
PR 10-JUN-1998; 98US-0088825P.
PR 10-JUN-1998; 98US-0088826P.
PR 11-JUN-1998; 98US-0088861P.
PR 11-JUN-1998; 98US-0088863P.
PR 11-JUN-1998; 98US-0088876P.
PR 12-JUN-1998; 98US-0089090P.
PR 12-JUN-1998; 98US-0089105P.
PR 16-JUN-1998; 98US-0089512P.
PR 16-JUN-1998; 98US-0089514P.
PR 17-JUN-1998; 98US-0089538P.
PR 17-JUN-1998; 98US-0089598P.
PR 17-JUN-1998; 98US-0089653P.
PR 18-JUN-1998; 98US-0089908P.
PR 19-JUN-1998; 98US-0089952P.
PR 22-JUN-1998; 98US-0090246P.
PR 22-JUN-1998; 98US-0090252P.
PR 22-JUN-1998; 98US-0090254P.
PR 24-JUN-1998; 98US-0090429P.
PR 24-JUN-1998; 98US-0090435P.
PR 24-JUN-1998; 98US-0090444P.
PR 24-JUN-1998; 98US-0090461P.
PR 24-JUN-1998; 98US-0090535P.
PR 24-JUN-1998; 98US-0090540P.
PR 25-JUN-1998; 98US-0090676P.
PR 25-JUN-1998; 98US-0090678P.
PR 25-JUN-1998; 98US-0090688P.
PR 25-JUN-1998; 98US-0090690P.
PR 25-JUN-1998; 98US-0090694P.
PR 25-JUN-1998; 98US-0090695P.
PR 25-JUN-1998; 98US-0090696P.
PR 26-JUN-1998; 98US-00105413.
PR 26-JUN-1998; 98US-0090862P.
PR 26-JUN-1998; 98US-0090863P.
PR 26-JUN-1998; 98US-0091010P.
PR 01-JUL-1998; 98US-0091359P.
PR 01-JUL-1998; 98US-0091544P.
PR 02-JUL-1998; 98US-0091478P.
PR 02-JUL-1998; 98US-0091486P.
PR 02-JUL-1998; 98US-0091626P.
PR 02-JUL-1998; 98US-0091628P.
PR 02-JUL-1998; 98US-0091632P.
PR 24-JUL-1998; 98US-0094006P.
PR 04-AUG-1998; 98US-0095282P.
PR 10-AUG-1998; 98US-0095998P.
PR 10-AUG-1998; 98US-0096012P.
PR 17-AUG-1998; 98US-0096757P.
PR 17-AUG-1998; 98US-0096766P.
PR 17-AUG-1998; 98US-0096867P.
PR 17-AUG-1998; 98US-0096891P.
PR 17-AUG-1998; 98US-0096897P.
PR 18-AUG-1998; 98US-0096949P.
PR 18-AUG-1998; 98US-0096959P.
PR 18-AUG-1998; 98US-0097022P.
PR 26-AUG-1998; 98US-0097952P.
PR 26-AUG-1998; 98US-0097954P.
PR 26-AUG-1998; 98US-0097955P.
PR 26-AUG-1998; 98US-0097971P.
PR 26-AUG-1998; 98US-0097974P.
PR 26-AUG-1998; 98US-0098014P.
PR 01-SEP-1998; 98US-0098716P.
PR 01-SEP-1998; 98US-0098723P.
PR 02-SEP-1998; 98US-0098803P.
PR 02-SEP-1998; 98US-0098821P.
PR 02-SEP-1998; 98US-0098843P.
PR 09-SEP-1998; 98US-0099602P.
PR 10-SEP-1998; 98US-0099741P.
PR 10-SEP-1998; 98US-0099754P.
PR 10-SEP-1998; 98US-0099763P.
PR 10-SEP-1998; 98US-0099812P.
PR 15-SEP-1998; 98US-0100388P.
PR 16-SEP-1998; 98US-0100662P.
PR 16-SEP-1998; 98US-0100664P.
PR 16-SEP-1998; 98US-0101751P.
PR 16-SEP-1998; 98WO-US019330.
PR 17-SEP-1998; 98US-0100683P.
PR 17-SEP-1998; 98US-0100684P.
PR 17-SEP-1998; 98US-0100919P.
PR 17-SEP-1998; 98US-0100930P.
PR 18-SEP-1998; 98US-0100849P.
PR 18-SEP-1998; 98US-0101014P.
PR 18-SEP-1998; 98US-0101068P.
PR 23-SEP-1998; 98US-0101471P.
PR 23-SEP-1998; 98US-0101472P.
PR 23-SEP-1998; 98US-0101475P.
PR 23-SEP-1998; 98US-0101477P.
PR 24-SEP-1998; 98US-0101738P.
PR 24-SEP-1998; 98US-0101739P.
PR 24-SEP-1998; 98US-0101743P.
PR 24-SEP-1998; 98US-0101922P.
PR 25-SEP-1998; 98US-0101786P.
PR 29-SEP-1998; 98US-0102207P.
PR 29-SEP-1998; 98US-0102240P.
PR 29-SEP-1998; 98US-0102330P.
PR 29-SEP-1998; 98US-0102331P.
PR 30-SEP-1998; 98US-0102487P.
PR 30-SEP-1998; 98US-0102570P.
PR 30-SEP-1998; 98US-0102571P.
PR 01-OCT-1998; 98US-0102684P.
PR 01-OCT-1998; 98US-0102687P.
PR 02-OCT-1998; 98US-0102965P.
PR 06-OCT-1998; 98US-0103258P.
PR 06-OCT-1998; 98US-0103449P.
PR 07-OCT-1998; 98US-00168978.

Query Match 98.7%; Score 1101; DB 6; Length 382;
Best Local Similarity 96.3%; Pred. No. 5.2e-103;
Matches 206; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

Qy 1 ATGRLLS--------GQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 52

```
Db         22 ATGRLLSASDLDLRGGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 81

Qy         53 ESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVD 112
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 ESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVD 141

Qy        113 EPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGE 172
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 EPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGE 201

Qy        173 ETELTTPVLPEETQEEDAKKTFKESREAALNLAY 206
              |||||||||||||||||||||||||||||||||
Db        202 ETELTTPVLPEETQEEDAKKTFKESREAALNLAY 235
```

RESULT 17
ABR66146
ID ABR66146 standard; protein; 382 AA.
XX
AC ABR66146;
XX
DT 05-AUG-2003 (first entry)
XX
DE Human secreted polypeptide PRO234, SEQ ID NO:20.
XX
KW Human; PRO; secreted protein; transmembrane protein;
KW extracellular domain; tumour necrosis factor-alpha; TNF-alpha;
KW chondrocyte; proliferation; differentiation; cartilage disorder;
KW bone disorder; arthritis; sports injury; cancer; tumour; diagnosis;
KW adrenal tumour; lung; colon; breast; prostate; kidney; rectum; cervix;
KW liver; drug screening; transgenic animal; genetic analysis;
KW antiarthritic; vulnerary; gene therapy.
XX
OS Homo sapiens.
XX
PN US2003027278-A1.
XX
PD 06-FEB-2003.
XX
PF 21-JUN-2002; 2002US-00176987.
XX
PR 18-SEP-1997; 97US-0059263P.
PR 18-SEP-1997; 97US-0059266P.
PR 17-OCT-1997; 97US-0062250P.
PR 21-OCT-1997; 97US-0063486P.
PR 24-OCT-1997; 97US-0063120P.
PR 24-OCT-1997; 97US-0063121P.
PR 28-OCT-1997; 97US-0063540P.
PR 28-OCT-1997; 97US-0063541P.
PR 28-OCT-1997; 97US-0063544P.
PR 28-OCT-1997; 97US-0063564P.
PR 29-OCT-1997; 97US-0063734P.
PR 31-OCT-1997; 97US-0063870P.
PR 31-OCT-1997; 97US-0064103P.
PR 13-NOV-1997; 97US-0065311P.
PR 21-NOV-1997; 97US-0066120P.
PR 24-NOV-1997; 97US-0066466P.
PR 24-NOV-1997; 97US-0066772P.
PR 11-DEC-1997; 97US-0069335P.
PR 12-DEC-1997; 97US-0069425P.
PR 17-DEC-1997; 97US-0069870P.
PR 18-DEC-1997; 97US-0068017P.
PR 10-MAR-1998; 98US-0077450P.
PR 11-MAR-1998; 98US-0077632P.
PR 11-MAR-1998; 98US-0077649P.
PR 20-MAR-1998; 98US-0078886P.
PR 20-MAR-1998; 98US-0078939P.
PR 27-MAR-1998; 98US-0079664P.
PR 27-MAR-1998; 98US-0079786P.
PR 31-MAR-1998; 98US-0080107P.
PR 31-MAR-1998; 98US-0080194P.
PR 01-APR-1998; 98US-0080327P.
PR 01-APR-1998; 98US-0080333P.
PR 08-APR-1998; 98US-0081049P.
PR 08-APR-1998; 98US-0081070P.
PR 09-APR-1998; 98US-0081195P.
PR 15-APR-1998; 98US-0081838P.
PR 21-APR-1998; 98US-0082568P.
PR 21-APR-1998; 98US-0082569P.
PR 22-APR-1998; 98US-0082704P.
PR 22-APR-1998; 98US-0082797P.
PR 28-APR-1998; 98US-0083322P.
PR 29-APR-1998; 98US-0083495P.
PR 29-APR-1998; 98US-0083496P.
PR 29-APR-1998; 98US-0083499P.
PR 29-APR-1998; 98US-0083559P.
PR 05-MAY-1998; 98US-0084366P.
PR 06-MAY-1998; 98US-0084414P.
PR 07-MAY-1998; 98US-0084639P.
PR 07-MAY-1998; 98US-0084640P.
PR 07-MAY-1998; 98US-0084643P.
PR 15-MAY-1998; 98US-0085579P.
PR 15-MAY-1998; 98US-0085580P.
PR 15-MAY-1998; 98US-0085582P.
PR 15-MAY-1998; 98US-0085700P.
PR 18-MAY-1998; 98US-0086023P.
PR 22-MAY-1998; 98US-0086392P.
PR 22-MAY-1998; 98US-0086486P.
PR 28-MAY-1998; 98US-0087098P.
PR 28-MAY-1998; 98US-0087208P.
PR 02-JUN-1998; 98US-0087609P.
PR 02-JUN-1998; 98US-0087759P.
PR 03-JUN-1998; 98US-0087827P.
PR 04-JUN-1998; 98US-0088025P.
PR 04-JUN-1998; 98US-0088028P.
PR 04-JUN-1998; 98US-0088029P.
PR 04-JUN-1998; 98US-0088033P.
PR 04-JUN-1998; 98US-0088326P.
PR 05-JUN-1998; 98US-0088167P.
PR 05-JUN-1998; 98US-0088202P.
PR 05-JUN-1998; 98US-0088212P.
PR 05-JUN-1998; 98US-0088217P.
PR 09-JUN-1998; 98US-0088655P.
PR 10-JUN-1998; 98US-0088722P.
PR 10-JUN-1998; 98US-0088738P.
PR 10-JUN-1998; 98US-0088740P.
PR 10-JUN-1998; 98US-0088811P.
PR 10-JUN-1998; 98US-0088824P.
PR 10-JUN-1998; 98US-0088825P.
PR 10-JUN-1998; 98US-0088826P.
PR 11-JUN-1998; 98US-0088861P.
PR 11-JUN-1998; 98US-0088863P.
PR 11-JUN-1998; 98US-0088876P.
PR 12-JUN-1998; 98US-0089090P.
PR 12-JUN-1998; 98US-0089105P.
PR 16-JUN-1998; 98US-0089512P.
PR 16-JUN-1998; 98US-0089514P.
PR 17-JUN-1998; 98US-0089538P.
PR 17-JUN-1998; 98US-0089598P.
PR 17-JUN-1998; 98US-0089653P.
PR 18-JUN-1998; 98US-0089908P.
PR 19-JUN-1998; 98US-0089952P.
PR 22-JUN-1998; 98US-0090246P.
PR 22-JUN-1998; 98US-0090252P.
PR 22-JUN-1998; 98US-0090254P.
PR 24-JUN-1998; 98US-0090429P.
PR 24-JUN-1998; 98US-0090435P.
PR 24-JUN-1998; 98US-0090444P.
PR 24-JUN-1998; 98US-0090461P.
PR 24-JUN-1998; 98US-0090535P.
PR 24-JUN-1998; 98US-0090540P.
PR 25-JUN-1998; 98US-0090676P.
PR 25-JUN-1998; 98US-0090678P.
PR 25-JUN-1998; 98US-0090688P.
PR 25-JUN-1998; 98US-0090690P.
PR 25-JUN-1998; 98US-0090694P.

PR 25-JUN-1998; 98US-0090695P.
PR 25-JUN-1998; 98US-0090696P.
PR 26-JUN-1998; 98US-00105413.
PR 26-JUN-1998; 98US-0090862P.
PR 26-JUN-1998; 98US-0090863P.
PR 26-JUN-1998; 98US-0091010P.
PR 01-JUL-1998; 98US-0091359P.
PR 01-JUL-1998; 98US-0091544P.
PR 02-JUL-1998; 98US-0091478P.
PR 02-JUL-1998; 98US-0091486P.
PR 02-JUL-1998; 98US-0091626P.
PR 02-JUL-1998; 98US-0091628P.
PR 02-JUL-1998; 98US-0091632P.
PR 24-JUL-1998; 98US-0094006P.
PR 04-AUG-1998; 98US-0095282P.
PR 10-AUG-1998; 98US-0095998P.
PR 10-AUG-1998; 98US-0096012P.
PR 17-AUG-1998; 98US-0096757P.
PR 17-AUG-1998; 98US-0096766P.
PR 17-AUG-1998; 98US-0096867P.
PR 17-AUG-1998; 98US-0096891P.
PR 17-AUG-1998; 98US-0096897P.
PR 18-AUG-1998; 98US-0096949P.
PR 18-AUG-1998; 98US-0096959P.
PR 18-AUG-1998; 98US-0097022P.
PR 26-AUG-1998; 98US-0097952P.
PR 26-AUG-1998; 98US-0097954P.
PR 26-AUG-1998; 98US-0097955P.
PR 26-AUG-1998; 98US-0097971P.
PR 26-AUG-1998; 98US-0097974P.
PR 26-AUG-1998; 98US-0098014P.
PR 01-SEP-1998; 98US-0098716P.
PR 01-SEP-1998; 98US-0098723P.
PR 02-SEP-1998; 98US-0098803P.
PR 02-SEP-1998; 98US-0098821P.
PR 02-SEP-1998; 98US-0098843P.
PR 09-SEP-1998; 98US-0099602P.
PR 10-SEP-1998; 98US-0099741P.
PR 10-SEP-1998; 98US-0099754P.
PR 10-SEP-1998; 98US-0099763P.
PR 10-SEP-1998; 98US-0099812P.
PR 15-SEP-1998; 98US-0100388P.
PR 16-SEP-1998; 98US-0100662P.
PR 16-SEP-1998; 98US-0100664P.
PR 16-SEP-1998; 98US-0101751P.
PR 16-SEP-1998; 98WO-US019330.
PR 17-SEP-1998; 98US-0100683P.
PR 17-SEP-1998; 98US-0100684P.
PR 17-SEP-1998; 98US-0100919P.
PR 17-SEP-1998; 98US-0100930P.
PR 18-SEP-1998; 98US-0100849P.
PR 18-SEP-1998; 98US-0101014P.
PR 18-SEP-1998; 98US-0101068P.
PR 23-SEP-1998; 98US-0101471P.
PR 23-SEP-1998; 98US-0101472P.
PR 23-SEP-1998; 98US-0101475P.
PR 23-SEP-1998; 98US-0101477P.
PR 24-SEP-1998; 98US-0101738P.
PR 24-SEP-1998; 98US-0101739P.
PR 24-SEP-1998; 98US-0101743P.
PR 24-SEP-1998; 98US-0101922P.
PR 25-SEP-1998; 98US-0101786P.
PR 29-SEP-1998; 98US-0102207P.
PR 29-SEP-1998; 98US-0102240P.
PR 29-SEP-1998; 98US-0102330P.
PR 29-SEP-1998; 98US-0102331P.
PR 30-SEP-1998; 98US-0102487P.
PR 30-SEP-1998; 98US-0102570P.
PR 30-SEP-1998; 98US-0102571P.
PR 01-OCT-1998; 98US-0102684P.
PR 01-OCT-1998; 98US-0102687P.

Query Match 98.7%; Score 1101; DB 6; Length 382;
Best Local Similarity 96.3%; Pred. No. 5.2e-103;
Matches 206; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

```
Qy          1 ATGRLLS--------GQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 52
              |||||||        ||||||||||||||||||||||||||||||||||||||||||||
Db         22 ATGRLLSASDLDLRGGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 81

Qy         53 ESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVD 112
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 ESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVD 141

Qy        113 EPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGE 172
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 EPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGE 201

Qy        173 ETELTTPVLPEETQEEDAKKTFKESREAALNLAY 206
              |||||||||||||||||||||||||||||||||
Db        202 ETELTTPVLPEETQEEDAKKTFKESREAALNLAY 235
```

RESULT 18
ABR65536
ID ABR65536 standard; protein; 382 AA.
XX
AC ABR65536;
XX
DT 05-AUG-2003 (first entry)
XX
DE Human secreted polypeptide PRO234, SEQ ID NO:20.
XX
KW Human; PRO; secreted protein; transmembrane protein;
KW extracellular domain; tumour necrosis factor-alpha; TNF-alpha;
KW chondrocyte; proliferation; differentiation; cartilage disorder;
KW bone disorder; arthritis; sports injury; cancer; tumour; diagnosis;
KW adrenal tumour; lung; colon; breast; prostate; kidney; rectum; cervix;
KW liver; drug screening; transgenic animal; genetic analysis;
KW antiarthritic; vulnerary; gene therapy.
XX
OS Homo sapiens.
XX
PN US2003036159-A1.
XX
PD 20-FEB-2003.
XX
PF 02-JUL-2002; 2002US-00188773.
XX
PR 18-SEP-1997; 97US-0059263P.
PR 18-SEP-1997; 97US-0059266P.
PR 17-OCT-1997; 97US-0062250P.
PR 21-OCT-1997; 97US-0063486P.
PR 24-OCT-1997; 97US-0063120P.
PR 24-OCT-1997; 97US-0063121P.
PR 28-OCT-1997; 97US-0063540P.
PR 28-OCT-1997; 97US-0063541P.
PR 28-OCT-1997; 97US-0063544P.
PR 28-OCT-1997; 97US-0063564P.
PR 29-OCT-1997; 97US-0063734P.
PR 31-OCT-1997; 97US-0063870P.
PR 31-OCT-1997; 97US-0064103P.
PR 13-NOV-1997; 97US-0065311P.
PR 21-NOV-1997; 97US-0066120P.
PR 24-NOV-1997; 97US-0066466P.
PR 24-NOV-1997; 97US-0066772P.
PR 11-DEC-1997; 97US-0069335P.
PR 12-DEC-1997; 97US-0069425P.
PR 17-DEC-1997; 97US-0069870P.
PR 18-DEC-1997; 97US-0068017P.
PR 10-MAR-1998; 98US-0077450P.
PR 11-MAR-1998; 98US-0077632P.
PR 11-MAR-1998; 98US-0077649P.
PR 20-MAR-1998; 98US-0078886P.
PR 20-MAR-1998; 98US-0078939P.
PR 27-MAR-1998; 98US-0079664P.

PR 27-MAR-1998; 98US-0079786P.
PR 31-MAR-1998; 98US-0080107P.
PR 31-MAR-1998; 98US-0080194P.
PR 01-APR-1998; 98US-0080327P.
PR 01-APR-1998; 98US-0080333P.
PR 08-APR-1998; 98US-0081049P.
PR 08-APR-1998; 98US-0081070P.
PR 09-APR-1998; 98US-0081195P.
PR 15-APR-1998; 98US-0081838P.
PR 21-APR-1998; 98US-0082568P.
PR 21-APR-1998; 98US-0082569P.
PR 22-APR-1998; 98US-0082704P.
PR 22-APR-1998; 98US-0082797P.
PR 28-APR-1998; 98US-0083322P.
PR 29-APR-1998; 98US-0083495P.
PR 29-APR-1998; 98US-0083496P.
PR 29-APR-1998; 98US-0083499P.
PR 29-APR-1998; 98US-0083559P.
PR 05-MAY-1998; 98US-0084366P.
PR 06-MAY-1998; 98US-0084414P.
PR 07-MAY-1998; 98US-0084639P.
PR 07-MAY-1998; 98US-0084640P.
PR 07-MAY-1998; 98US-0084643P.
PR 15-MAY-1998; 98US-0085579P.
PR 15-MAY-1998; 98US-0085580P.
PR 15-MAY-1998; 98US-0085582P.
PR 15-MAY-1998; 98US-0085700P.
PR 18-MAY-1998; 98US-0086023P.
PR 22-MAY-1998; 98US-0086392P.
PR 22-MAY-1998; 98US-0086486P.
PR 28-MAY-1998; 98US-0087098P.
PR 28-MAY-1998; 98US-0087208P.
PR 02-JUN-1998; 98US-0087609P.
PR 02-JUN-1998; 98US-0087759P.
PR 03-JUN-1998; 98US-0087827P.
PR 04-JUN-1998; 98US-0088025P.
PR 04-JUN-1998; 98US-0088028P.
PR 04-JUN-1998; 98US-0088029P.
PR 04-JUN-1998; 98US-0088033P.
PR 04-JUN-1998; 98US-0088326P.
PR 05-JUN-1998; 98US-0088167P.
PR 05-JUN-1998; 98US-0088202P.
PR 05-JUN-1998; 98US-0088212P.
PR 05-JUN-1998; 98US-0088217P.
PR 09-JUN-1998; 98US-0088655P.
PR 10-JUN-1998; 98US-0088722P.
PR 10-JUN-1998; 98US-0088738P.
PR 10-JUN-1998; 98US-0088740P.
PR 10-JUN-1998; 98US-0088811P.
PR 10-JUN-1998; 98US-0088824P.
PR 10-JUN-1998; 98US-0088825P.
PR 10-JUN-1998; 98US-0088826P.
PR 11-JUN-1998; 98US-0088861P.
PR 11-JUN-1998; 98US-0088863P.
PR 11-JUN-1998; 98US-0088876P.
PR 12-JUN-1998; 98US-0089090P.
PR 12-JUN-1998; 98US-0089105P.
PR 16-JUN-1998; 98US-0089512P.
PR 16-JUN-1998; 98US-0089514P.
PR 17-JUN-1998; 98US-0089538P.
PR 17-JUN-1998; 98US-0089598P.
PR 17-JUN-1998; 98US-0089653P.
PR 18-JUN-1998; 98US-0089908P.
PR 19-JUN-1998; 98US-0089952P.
PR 22-JUN-1998; 98US-0090246P.
PR 22-JUN-1998; 98US-0090252P.
PR 22-JUN-1998; 98US-0090254P.
PR 24-JUN-1998; 98US-0090429P.
PR 24-JUN-1998; 98US-0090435P.
PR 24-JUN-1998; 98US-0090444P.
PR 24-JUN-1998; 98US-0090461P.
PR 24-JUN-1998; 98US-0090535P.
PR 24-JUN-1998; 98US-0090540P.
PR 25-JUN-1998; 98US-0090676P.
PR 25-JUN-1998; 98US-0090678P.
PR 25-JUN-1998; 98US-0090688P.
PR 25-JUN-1998; 98US-0090690P.
PR 25-JUN-1998; 98US-0090694P.
PR 25-JUN-1998; 98US-0090695P.
PR 25-JUN-1998; 98US-0090696P.
PR 26-JUN-1998; 98US-00105413.
PR 26-JUN-1998; 98US-0090862P.
PR 26-JUN-1998; 98US-0090863P.
PR 26-JUN-1998; 98US-0091010P.
PR 01-JUL-1998; 98US-0091359P.
PR 01-JUL-1998; 98US-0091544P.
PR 02-JUL-1998; 98US-0091478P.
PR 02-JUL-1998; 98US-0091486P.
PR 02-JUL-1998; 98US-0091626P.
PR 02-JUL-1998; 98US-0091628P.
PR 02-JUL-1998; 98US-0091632P.
PR 24-JUL-1998; 98US-0094006P.
PR 04-AUG-1998; 98US-0095282P.
PR 10-AUG-1998; 98US-0095998P.
PR 10-AUG-1998; 98US-0096012P.
PR 17-AUG-1998; 98US-0096757P.
PR 17-AUG-1998; 98US-0096766P.
PR 17-AUG-1998; 98US-0096867P.
PR 17-AUG-1998; 98US-0096891P.
PR 17-AUG-1998; 98US-0096897P.
PR 18-AUG-1998; 98US-0096949P.
PR 18-AUG-1998; 98US-0096959P.
PR 18-AUG-1998; 98US-0097022P.
PR 26-AUG-1998; 98US-0097952P.
PR 26-AUG-1998; 98US-0097954P.
PR 26-AUG-1998; 98US-0097955P.
PR 26-AUG-1998; 98US-0097971P.
PR 26-AUG-1998; 98US-0097974P.
PR 26-AUG-1998; 98US-0098014P.
PR 01-SEP-1998; 98US-0098716P.
PR 01-SEP-1998; 98US-0098723P.
PR 02-SEP-1998; 98US-0098803P.
PR 02-SEP-1998; 98US-0098821P.
PR 02-SEP-1998; 98US-0098843P.
PR 09-SEP-1998; 98US-0099602P.
PR 10-SEP-1998; 98US-0099741P.
PR 10-SEP-1998; 98US-0099754P.
PR 10-SEP-1998; 98US-0099763P.
PR 10-SEP-1998; 98US-0099812P.
PR 15-SEP-1998; 98US-0100388P.
PR 16-SEP-1998; 98US-0100662P.
PR 16-SEP-1998; 98US-0100664P.
PR 16-SEP-1998; 98US-0101751P.
PR 16-SEP-1998; 98WO-US019330.
PR 17-SEP-1998; 98US-0100683P.
PR 17-SEP-1998; 98US-0100684P.
PR 17-SEP-1998; 98US-0100919P.
PR 17-SEP-1998; 98US-0100930P.
PR 18-SEP-1998; 98US-0100849P.
PR 18-SEP-1998; 98US-0101014P.
PR 18-SEP-1998; 98US-0101068P.
PR 23-SEP-1998; 98US-0101471P.
PR 23-SEP-1998; 98US-0101472P.
PR 23-SEP-1998; 98US-0101475P.
PR 23-SEP-1998; 98US-0101477P.
PR 24-SEP-1998; 98US-0101738P.
PR 24-SEP-1998; 98US-0101739P.
PR 24-SEP-1998; 98US-0101743P.
PR 24-SEP-1998; 98US-0101922P.
PR 25-SEP-1998; 98US-0101786P.
PR 29-SEP-1998; 98US-0102207P.
PR 29-SEP-1998; 98US-0102240P.
PR 29-SEP-1998; 98US-0102330P.
PR 29-SEP-1998; 98US-0102331P.
PR 30-SEP-1998; 98US-0102487P.
PR 30-SEP-1998; 98US-0102570P.

PR 30-SEP-1998; 98US-0102571P.
PR 01-OCT-1998; 98US-0102684P.
PR 01-OCT-1998; 98US-0102687P.

Query Match 98.7%; Score 1101; DB 6; Length 382;
Best Local Similarity 96.3%; Pred. No. 5.2e-103;
Matches 206; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

```
Qy          1 ATGRLLS--------GQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 52
              |||||||        ||||||||||||||||||||||||||||||||||||||||||||
Db         22 ATGRLLSASDLDLRGGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 81

Qy         53 ESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVD 112
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 ESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVD 141

Qy        113 EPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGE 172
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 EPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGE 201

Qy        173 ETELTTPVLPEETQEEDAKKTFKESREAALNLAY 206
              |||||||||||||||||||||||||||||||||
Db        202 ETELTTPVLPEETQEEDAKKTFKESREAALNLAY 235
```

RESULT 19
ABU99476
ID ABU99476 standard; protein; 382 AA.
XX
AC ABU99476;
XX
DT 09-AUG-2003 (first entry)
XX
DE Human secreted/transmembrane protein (PRO) #10.
XX
KW Human; secreted and transmembrane protein; PRO; TNF-alpha;
KW tumour necrosis factor alpha; chondrocyte cell; tumour; gene therapy;
KW tissue typing.
XX
OS Homo sapiens.
XX
PN US2003040070-A1.
XX
PD 27-FEB-2003.
XX
PF 27-JUN-2002; 2002US-00184627.
XX
PR 18-SEP-1997; 97US-0059263P.
PR 18-SEP-1997; 97US-0059266P.
PR 17-OCT-1997; 97US-0062250P.
PR 21-OCT-1997; 97US-0063486P.
PR 24-OCT-1997; 97US-0063120P.
PR 24-OCT-1997; 97US-0063121P.
PR 28-OCT-1997; 97US-0063540P.
PR 28-OCT-1997; 97US-0063541P.
PR 28-OCT-1997; 97US-0063544P.
PR 28-OCT-1997; 97US-0063564P.
PR 29-OCT-1997; 97US-0063734P.
PR 31-OCT-1997; 97US-0063870P.
PR 31-OCT-1997; 97US-0064103P.
PR 13-NOV-1997; 97US-0065311P.
PR 21-NOV-1997; 97US-0066120P.
PR 24-NOV-1997; 97US-0066466P.
PR 24-NOV-1997; 97US-0066772P.
PR 11-DEC-1997; 97US-0069335P.
PR 12-DEC-1997; 97US-0069425P.
PR 17-DEC-1997; 97US-0069870P.
PR 18-DEC-1997; 97US-0068017P.
PR 10-MAR-1998; 98US-0077450P.
PR 11-MAR-1998; 98US-0077632P.
PR 11-MAR-1998; 98US-0077649P.
PR 20-MAR-1998; 98US-0078886P.
PR 20-MAR-1998; 98US-0078939P.
PR 27-MAR-1998; 98US-0079664P.
PR 27-MAR-1998; 98US-0079786P.
PR 31-MAR-1998; 98US-0080107P.
PR 31-MAR-1998; 98US-0080194P.
PR 01-APR-1998; 98US-0080327P.
PR 01-APR-1998; 98US-0080333P.
PR 08-APR-1998; 98US-0081049P.
PR 08-APR-1998; 98US-0081070P.
PR 09-APR-1998; 98US-0081195P.
PR 15-APR-1998; 98US-0081838P.
PR 21-APR-1998; 98US-0082568P.
PR 21-APR-1998; 98US-0082569P.
PR 22-APR-1998; 98US-0082704P.
PR 22-APR-1998; 98US-0082797P.
PR 28-APR-1998; 98US-0083322P.
PR 29-APR-1998; 98US-0083495P.
PR 29-APR-1998; 98US-0083496P.
PR 29-APR-1998; 98US-0083499P.
PR 29-APR-1998; 98US-0083559P.
PR 05-MAY-1998; 98US-0084366P.
PR 06-MAY-1998; 98US-0084414P.
PR 07-MAY-1998; 98US-0084639P.
PR 07-MAY-1998; 98US-0084640P.
PR 07-MAY-1998; 98US-0084643P.
PR 15-MAY-1998; 98US-0085579P.
PR 15-MAY-1998; 98US-0085580P.
PR 15-MAY-1998; 98US-0085582P.
PR 15-MAY-1998; 98US-0085700P.
PR 18-MAY-1998; 98US-0086023P.
PR 22-MAY-1998; 98US-0086392P.
PR 22-MAY-1998; 98US-0086486P.
PR 28-MAY-1998; 98US-0087098P.
PR 28-MAY-1998; 98US-0087208P.
PR 02-JUN-1998; 98US-0087609P.
PR 02-JUN-1998; 98US-0087759P.
PR 03-JUN-1998; 98US-0087827P.
PR 04-JUN-1998; 98US-0088025P.
PR 04-JUN-1998; 98US-0088028P.
PR 04-JUN-1998; 98US-0088029P.
PR 04-JUN-1998; 98US-0088033P.
PR 04-JUN-1998; 98US-0088326P.
PR 05-JUN-1998; 98US-0088167P.
PR 05-JUN-1998; 98US-0088202P.
PR 05-JUN-1998; 98US-0088212P.
PR 05-JUN-1998; 98US-0088217P.
PR 09-JUN-1998; 98US-0088655P.
PR 10-JUN-1998; 98US-0088722P.
PR 10-JUN-1998; 98US-0088738P.
PR 10-JUN-1998; 98US-0088740P.
PR 10-JUN-1998; 98US-0088811P.
PR 10-JUN-1998; 98US-0088824P.
PR 10-JUN-1998; 98US-0088825P.
PR 10-JUN-1998; 98US-0088826P.
PR 11-JUN-1998; 98US-0088861P.
PR 11-JUN-1998; 98US-0088863P.
PR 11-JUN-1998; 98US-0088876P.
PR 12-JUN-1998; 98US-0089090P.
PR 12-JUN-1998; 98US-0089105P.
PR 16-JUN-1998; 98US-0089512P.
PR 16-JUN-1998; 98US-0089514P.
PR 17-JUN-1998; 98US-0089538P.
PR 17-JUN-1998; 98US-0089598P.
PR 17-JUN-1998; 98US-0089653P.
PR 18-JUN-1998; 98US-0089908P.
PR 19-JUN-1998; 98US-0089952P.
PR 22-JUN-1998; 98US-0090246P.
PR 22-JUN-1998; 98US-0090252P.
PR 22-JUN-1998; 98US-0090254P.
PR 24-JUN-1998; 98US-0090429P.
PR 24-JUN-1998; 98US-0090435P.
PR 24-JUN-1998; 98US-0090444P.
PR 24-JUN-1998; 98US-0090461P.
PR 24-JUN-1998; 98US-0090535P.

PR 24-JUN-1998; 98US-0090540P.
PR 25-JUN-1998; 98US-0090676P.
PR 25-JUN-1998; 98US-0090678P.
PR 25-JUN-1998; 98US-0090688P.
PR 25-JUN-1998; 98US-0090690P.
PR 25-JUN-1998; 98US-0090694P.
PR 25-JUN-1998; 98US-0090695P.
PR 25-JUN-1998; 98US-0090696P.
PR 26-JUN-1998; 98US-00105413.
PR 26-JUN-1998; 98US-0090862P.
PR 26-JUN-1998; 98US-0090863P.
PR 26-JUN-1998; 98US-0091010P.
PR 01-JUL-1998; 98US-0091359P.
PR 01-JUL-1998; 98US-0091544P.
PR 02-JUL-1998; 98US-0091478P.
PR 02-JUL-1998; 98US-0091486P.
PR 02-JUL-1998; 98US-0091626P.
PR 02-JUL-1998; 98US-0091628P.
PR 02-JUL-1998; 98US-0091632P.
PR 24-JUL-1998; 98US-0094006P.
PR 04-AUG-1998; 98US-0095282P.
PR 10-AUG-1998; 98US-0095998P.
PR 10-AUG-1998; 98US-0096012P.
PR 17-AUG-1998; 98US-0096757P.
PR 17-AUG-1998; 98US-0096766P.
PR 17-AUG-1998; 98US-0096867P.
PR 17-AUG-1998; 98US-0096891P.
PR 17-AUG-1998; 98US-0096897P.
PR 18-AUG-1998; 98US-0096949P.
PR 18-AUG-1998; 98US-0096959P.
PR 18-AUG-1998; 98US-0097022P.
PR 26-AUG-1998; 98US-0097952P.
PR 26-AUG-1998; 98US-0097954P.
PR 26-AUG-1998; 98US-0097955P.
PR 26-AUG-1998; 98US-0097971P.
PR 26-AUG-1998; 98US-0097974P.
PR 26-AUG-1998; 98US-0098014P.
PR 01-SEP-1998; 98US-0098716P.
PR 01-SEP-1998; 98US-0098723P.
PR 02-SEP-1998; 98US-0098803P.
PR 02-SEP-1998; 98US-0098821P.
PR 02-SEP-1998; 98US-0098843P.
PR 09-SEP-1998; 98US-0099602P.
PR 10-SEP-1998; 98US-0099741P.
PR 10-SEP-1998; 98US-0099754P.
PR 10-SEP-1998; 98US-0099763P.
PR 10-SEP-1998; 98US-0099812P.
PR 15-SEP-1998; 98US-0100388P.
PR 16-SEP-1998; 98US-0100662P.
PR 16-SEP-1998; 98US-0100664P.
PR 16-SEP-1998; 98US-0101751P.
PR 16-SEP-1998; 98WO-US019330.
PR 17-SEP-1998; 98US-0100683P.
PR 17-SEP-1998; 98US-0100684P.
PR 17-SEP-1998; 98US-0100919P.
PR 17-SEP-1998; 98US-0100930P.
PR 18-SEP-1998; 98US-0100849P.
PR 18-SEP-1998; 98US-0101014P.
PR 18-SEP-1998; 98US-0101068P.
PR 23-SEP-1998; 98US-0101471P.
PR 23-SEP-1998; 98US-0101472P.
PR 23-SEP-1998; 98US-0101475P.
PR 23-SEP-1998; 98US-0101477P.
PR 24-SEP-1998; 98US-0101738P.
PR 24-SEP-1998; 98US-0101739P.
PR 24-SEP-1998; 98US-0101743P.
PR 24-SEP-1998; 98US-0101922P.
PR 25-SEP-1998; 98US-0101786P.
PR 29-SEP-1998; 98US-0102207P.
PR 29-SEP-1998; 98US-0102240P.
PR 29-SEP-1998; 98US-0102330P.
PR 29-SEP-1998; 98US-0102331P.
PR 30-SEP-1998; 98US-0102487P.
PR 30-SEP-1998; 98US-0102570P.
PR 30-SEP-1998; 98US-0102571P.
PR 01-OCT-1998; 98US-0102684P.
PR 01-OCT-1998; 98US-0102687P.
PR 02-OCT-1998; 98US-0102965P.
PR 06-OCT-1998; 98US-0103258P.
PR 06-OCT-1998; 98US-0103449P.
PR 07-OCT-1998; 98US-00168978.

Query Match 98.7%; Score 1101; DB 6; Length 382;
Best Local Similarity 96.3%; Pred. No. 5.2e-103;
Matches 206; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

```
Qy          1 ATGRLLS--------GQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 52
              |||||||        ||||||||||||||||||||||||||||||||||||||||||||
Db         22 ATGRLLSASDLDLRGGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 81

Qy         53 ESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVD 112
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 ESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVD 141

Qy        113 EPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGE 172
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 EPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGE 201

Qy        173 ETELTTPVLPEETQEEDAKKTFKESREAALNLAY 206
              |||||||||||||||||||||||||||||||||
Db        202 ETELTTPVLPEETQEEDAKKTFKESREAALNLAY 235
```

RESULT 20
ABU82715
ID ABU82715 standard; protein; 382 AA.
XX
AC ABU82715;
XX
DT 27-JUN-2003 (first entry)
XX
DE Human PRO polypeptide #10.
XX
KW Human; PRO polypeptide; secreted and transmembrane protein; tumour;
KW chromosome mapping; gene mapping; cytostatic.
XX
OS Homo sapiens.
XX
PN US2003032113-A1.
XX
PD 13-FEB-2003.
XX
PF 20-JUN-2002; 2002US-00176911.
XX
PR 18-SEP-1997; 97US-0059263P.
PR 18-SEP-1997; 97US-0059266P.
PR 17-OCT-1997; 97US-0062250P.
PR 21-OCT-1997; 97US-0063486P.
PR 24-OCT-1997; 97US-0063120P.
PR 24-OCT-1997; 97US-0063121P.
PR 28-OCT-1997; 97US-0063540P.
PR 28-OCT-1997; 97US-0063541P.
PR 28-OCT-1997; 97US-0063544P.
PR 28-OCT-1997; 97US-0063564P.
PR 29-OCT-1997; 97US-0063734P.
PR 31-OCT-1997; 97US-0063870P.
PR 31-OCT-1997; 97US-0064103P.
PR 13-NOV-1997; 97US-0065311P.
PR 21-NOV-1997; 97US-0066120P.
PR 24-NOV-1997; 97US-0066466P.
PR 24-NOV-1997; 97US-0066772P.
PR 11-DEC-1997; 97US-0069335P.
PR 12-DEC-1997; 97US-0069425P.
PR 17-DEC-1997; 97US-0069870P.
PR 18-DEC-1997; 97US-0068017P.
PR 10-MAR-1998; 98US-0077450P.

PR 11-MAR-1998; 98US-0077632P.
PR 11-MAR-1998; 98US-0077649P.
PR 20-MAR-1998; 98US-0078886P.
PR 20-MAR-1998; 98US-0078939P.
PR 27-MAR-1998; 98US-0079664P.
PR 27-MAR-1998; 98US-0079786P.
PR 31-MAR-1998; 98US-0080107P.
PR 31-MAR-1998; 98US-0080194P.
PR 01-APR-1998; 98US-0080327P.
PR 01-APR-1998; 98US-0080333P.
PR 08-APR-1998; 98US-0081049P.
PR 08-APR-1998; 98US-0081070P.
PR 09-APR-1998; 98US-0081195P.
PR 15-APR-1998; 98US-0081838P.
PR 21-APR-1998; 98US-0082568P.
PR 21-APR-1998; 98US-0082569P.
PR 22-APR-1998; 98US-0082704P.
PR 22-APR-1998; 98US-0082797P.
PR 28-APR-1998; 98US-0083322P.
PR 29-APR-1998; 98US-0083495P.
PR 29-APR-1998; 98US-0083496P.
PR 29-APR-1998; 98US-0083499P.
PR 29-APR-1998; 98US-0083559P.
PR 05-MAY-1998; 98US-0084366P.
PR 06-MAY-1998; 98US-0084414P.
PR 07-MAY-1998; 98US-0084639P.
PR 07-MAY-1998; 98US-0084640P.
PR 07-MAY-1998; 98US-0084643P.
PR 15-MAY-1998; 98US-0085579P.
PR 15-MAY-1998; 98US-0085580P.
PR 15-MAY-1998; 98US-0085582P.
PR 15-MAY-1998; 98US-0085700P.
PR 18-MAY-1998; 98US-0086023P.
PR 22-MAY-1998; 98US-0086392P.
PR 22-MAY-1998; 98US-0086486P.
PR 28-MAY-1998; 98US-0087098P.
PR 28-MAY-1998; 98US-0087208P.
PR 02-JUN-1998; 98US-0087609P.
PR 02-JUN-1998; 98US-0087759P.
PR 03-JUN-1998; 98US-0087827P.
PR 04-JUN-1998; 98US-0088025P.
PR 04-JUN-1998; 98US-0088028P.
PR 04-JUN-1998; 98US-0088029P.
PR 04-JUN-1998; 98US-0088033P.
PR 04-JUN-1998; 98US-0088326P.
PR 05-JUN-1998; 98US-0088167P.
PR 05-JUN-1998; 98US-0088202P.
PR 05-JUN-1998; 98US-0088212P.
PR 05-JUN-1998; 98US-0088217P.
PR 09-JUN-1998; 98US-0088655P.
PR 10-JUN-1998; 98US-0088722P.
PR 10-JUN-1998; 98US-0088738P.
PR 10-JUN-1998; 98US-0088740P.
PR 10-JUN-1998; 98US-0088811P.
PR 10-JUN-1998; 98US-0088824P.
PR 10-JUN-1998; 98US-0088825P.
PR 10-JUN-1998; 98US-0088826P.
PR 11-JUN-1998; 98US-0088861P.
PR 11-JUN-1998; 98US-0088863P.
PR 11-JUN-1998; 98US-0088876P.
PR 12-JUN-1998; 98US-0089090P.
PR 12-JUN-1998; 98US-0089105P.
PR 16-JUN-1998; 98US-0089512P.
PR 16-JUN-1998; 98US-0089514P.
PR 17-JUN-1998; 98US-0089538P.
PR 17-JUN-1998; 98US-0089598P.
PR 17-JUN-1998; 98US-0089653P.
PR 18-JUN-1998; 98US-0089908P.
PR 19-JUN-1998; 98US-0089952P.
PR 22-JUN-1998; 98US-0090246P.
PR 22-JUN-1998; 98US-0090252P.
PR 22-JUN-1998; 98US-0090254P.
PR 24-JUN-1998; 98US-0090429P.
PR 24-JUN-1998; 98US-0090435P.
PR 24-JUN-1998; 98US-0090444P.
PR 24-JUN-1998; 98US-0090461P.
PR 24-JUN-1998; 98US-0090535P.
PR 24-JUN-1998; 98US-0090540P.
PR 25-JUN-1998; 98US-0090676P.
PR 25-JUN-1998; 98US-0090678P.
PR 25-JUN-1998; 98US-0090688P.
PR 25-JUN-1998; 98US-0090690P.
PR 25-JUN-1998; 98US-0090694P.
PR 25-JUN-1998; 98US-0090695P.
PR 25-JUN-1998; 98US-0090696P.
PR 26-JUN-1998; 98US-00105413.
PR 26-JUN-1998; 98US-0090862P.
PR 26-JUN-1998; 98US-0090863P.
PR 26-JUN-1998; 98US-0091010P.
PR 01-JUL-1998; 98US-0091359P.
PR 01-JUL-1998; 98US-0091544P.
PR 02-JUL-1998; 98US-0091478P.
PR 02-JUL-1998; 98US-0091486P.
PR 02-JUL-1998; 98US-0091626P.
PR 02-JUL-1998; 98US-0091628P.
PR 02-JUL-1998; 98US-0091632P.
PR 24-JUL-1998; 98US-0094006P.
PR 04-AUG-1998; 98US-0095282P.
PR 10-AUG-1998; 98US-0095998P.
PR 10-AUG-1998; 98US-0096012P.
PR 17-AUG-1998; 98US-0096757P.
PR 17-AUG-1998; 98US-0096766P.
PR 17-AUG-1998; 98US-0096867P.
PR 17-AUG-1998; 98US-0096891P.
PR 17-AUG-1998; 98US-0096897P.
PR 18-AUG-1998; 98US-0096949P.
PR 18-AUG-1998; 98US-0096959P.
PR 18-AUG-1998; 98US-0097022P.
PR 26-AUG-1998; 98US-0097952P.
PR 26-AUG-1998; 98US-0097954P.
PR 26-AUG-1998; 98US-0097955P.
PR 26-AUG-1998; 98US-0097971P.
PR 26-AUG-1998; 98US-0097974P.
PR 26-AUG-1998; 98US-0098014P.
PR 01-SEP-1998; 98US-0098716P.
PR 01-SEP-1998; 98US-0098723P.
PR 02-SEP-1998; 98US-0098803P.
PR 02-SEP-1998; 98US-0098821P.
PR 02-SEP-1998; 98US-0098843P.
PR 09-SEP-1998; 98US-0099602P.
PR 10-SEP-1998; 98US-0099741P.
PR 10-SEP-1998; 98US-0099754P.
PR 10-SEP-1998; 98US-0099763P.
PR 10-SEP-1998; 98US-0099812P.
PR 15-SEP-1998; 98US-0100388P.
PR 16-SEP-1998; 98US-0100662P.
PR 16-SEP-1998; 98US-0100664P.
PR 16-SEP-1998; 98US-0101751P.
PR 16-SEP-1998; 98WO-US019330.
PR 17-SEP-1998; 98US-0100683P.
PR 17-SEP-1998; 98US-0100684P.
PR 17-SEP-1998; 98US-0100919P.
PR 17-SEP-1998; 98US-0100930P.
PR 18-SEP-1998; 98US-0100849P.
PR 18-SEP-1998; 98US-0101014P.
PR 18-SEP-1998; 98US-0101068P.
PR 23-SEP-1998; 98US-0101471P.
PR 23-SEP-1998; 98US-0101472P.
PR 23-SEP-1998; 98US-0101475P.
PR 23-SEP-1998; 98US-0101477P.
PR 24-SEP-1998; 98US-0101738P.
PR 24-SEP-1998; 98US-0101739P.
PR 24-SEP-1998; 98US-0101743P.
PR 24-SEP-1998; 98US-0101922P.
PR 25-SEP-1998; 98US-0101786P.
PR 29-SEP-1998; 98US-0102207P.

PR 29-SEP-1998; 98US-0102240P.
PR 29-SEP-1998; 98US-0102330P.
PR 29-SEP-1998; 98US-0102331P.
PR 30-SEP-1998; 98US-0102487P.
PR 30-SEP-1998; 98US-0102570P.
PR 30-SEP-1998; 98US-0102571P.
PR 01-OCT-1998; 98US-0102684P.
PR 01-OCT-1998; 98US-0102687P.
PR 02-OCT-1998; 98US-0102965P.
PR 06-OCT-1998; 98US-0103258P.
PR 06-OCT-1998; 98US-0103449P.
PR 07-OCT-1998; 98US-00168978.
PR 07-OCT-1998; 98US-0103395P.

Query Match 98.7%; Score 1101; DB 6; Length 382;
Best Local Similarity 96.3%; Pred. No. 5.2e-103;
Matches 206; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

```
Qy          1 ATGRLLS--------GQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 52
              |||||||        ||||||||||||||||||||||||||||||||||||||||||||
Db         22 ATGRLLSASDLDLRGGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 81

Qy         53 ESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVD 112
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 ESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVD 141

Qy        113 EPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGE 172
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 EPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGE 201

Qy        173 ETELTTPVLPEETQEEDAKKTFKESREAALNLAY 206
              |||||||||||||||||||||||||||||||||
Db        202 ETELTTPVLPEETQEEDAKKTFKESREAALNLAY 235
```

RESULT 21
ABU89836
ID ABU89836 standard; protein; 382 AA.
XX
AC ABU89836;
XX
DT 11-AUG-2003 (first entry)
XX
DE Novel human secreted and transmembrane protein PRO234.
XX
KW Human; gene therapy; tissue typing; tumour; chondrocyte proliferation;
KW chondrocyte differentiation; tumour necrosis factor-alpha release;
KW affinity purification.
XX
OS Homo sapiens.
XX
PN US2003036147-A1.
XX
PD 20-FEB-2003.
XX
PF 02-JUL-2002; 2002US-00187741.
XX
PR 18-SEP-1997; 97US-0059263P.
PR 18-SEP-1997; 97US-0059266P.
PR 17-OCT-1997; 97US-0062250P.
PR 21-OCT-1997; 97US-0063486P.
PR 24-OCT-1997; 97US-0063120P.
PR 24-OCT-1997; 97US-0063121P.
PR 28-OCT-1997; 97US-0063540P.
PR 28-OCT-1997; 97US-0063541P.
PR 28-OCT-1997; 97US-0063544P.
PR 28-OCT-1997; 97US-0063564P.
PR 29-OCT-1997; 97US-0063734P.
PR 31-OCT-1997; 97US-0063870P.
PR 31-OCT-1997; 97US-0064103P.
PR 13-NOV-1997; 97US-0065311P.
PR 21-NOV-1997; 97US-0066120P.
PR 24-NOV-1997; 97US-0066466P.
PR 24-NOV-1997; 97US-0066772P.
PR 11-DEC-1997; 97US-0069335P.
PR 12-DEC-1997; 97US-0069425P.
PR 17-DEC-1997; 97US-0069870P.
PR 18-DEC-1997; 97US-0068017P.
PR 10-MAR-1998; 98US-0077450P.
PR 11-MAR-1998; 98US-0077632P.
PR 11-MAR-1998; 98US-0077649P.
PR 20-MAR-1998; 98US-0078886P.
PR 20-MAR-1998; 98US-0078939P.
PR 27-MAR-1998; 98US-0079664P.
PR 27-MAR-1998; 98US-0079786P.
PR 31-MAR-1998; 98US-0080107P.
PR 31-MAR-1998; 98US-0080194P.
PR 01-APR-1998; 98US-0080327P.
PR 01-APR-1998; 98US-0080333P.
PR 08-APR-1998; 98US-0081049P.
PR 08-APR-1998; 98US-0081070P.
PR 09-APR-1998; 98US-0081195P.
PR 15-APR-1998; 98US-0081838P.
PR 21-APR-1998; 98US-0082568P.
PR 21-APR-1998; 98US-0082569P.
PR 22-APR-1998; 98US-0082704P.
PR 22-APR-1998; 98US-0082797P.
PR 28-APR-1998; 98US-0083322P.
PR 29-APR-1998; 98US-0083495P.
PR 29-APR-1998; 98US-0083496P.
PR 29-APR-1998; 98US-0083499P.
PR 29-APR-1998; 98US-0083559P.
PR 05-MAY-1998; 98US-0084366P.
PR 06-MAY-1998; 98US-0084414P.
PR 07-MAY-1998; 98US-0084639P.
PR 07-MAY-1998; 98US-0084640P.
PR 07-MAY-1998; 98US-0084643P.
PR 15-MAY-1998; 98US-0085579P.
PR 15-MAY-1998; 98US-0085580P.
PR 15-MAY-1998; 98US-0085582P.
PR 15-MAY-1998; 98US-0085700P.
PR 18-MAY-1998; 98US-0086023P.
PR 22-MAY-1998; 98US-0086392P.
PR 22-MAY-1998; 98US-0086486P.
PR 28-MAY-1998; 98US-0087098P.
PR 28-MAY-1998; 98US-0087208P.
PR 02-JUN-1998; 98US-0087609P.
PR 02-JUN-1998; 98US-0087759P.
PR 03-JUN-1998; 98US-0087827P.
PR 04-JUN-1998; 98US-0088025P.
PR 04-JUN-1998; 98US-0088028P.
PR 04-JUN-1998; 98US-0088029P.
PR 04-JUN-1998; 98US-0088033P.
PR 04-JUN-1998; 98US-0088326P.
PR 05-JUN-1998; 98US-0088167P.
PR 05-JUN-1998; 98US-0088202P.
PR 05-JUN-1998; 98US-0088212P.
PR 05-JUN-1998; 98US-0088217P.
PR 09-JUN-1998; 98US-0088655P.
PR 10-JUN-1998; 98US-0088722P.
PR 10-JUN-1998; 98US-0088738P.
PR 10-JUN-1998; 98US-0088740P.
PR 10-JUN-1998; 98US-0088811P.
PR 10-JUN-1998; 98US-0088824P.
PR 10-JUN-1998; 98US-0088825P.
PR 10-JUN-1998; 98US-0088826P.
PR 11-JUN-1998; 98US-0088861P.
PR 11-JUN-1998; 98US-0088863P.
PR 11-JUN-1998; 98US-0088876P.
PR 12-JUN-1998; 98US-0089090P.
PR 12-JUN-1998; 98US-0089105P.
PR 16-JUN-1998; 98US-0089512P.
PR 16-JUN-1998; 98US-0089514P.
PR 17-JUN-1998; 98US-0089538P.
PR 17-JUN-1998; 98US-0089598P.
PR 17-JUN-1998; 98US-0089653P.

PR 18-JUN-1998; 98US-0089908P.
PR 19-JUN-1998; 98US-0089952P.
PR 22-JUN-1998; 98US-0090246P.
PR 22-JUN-1998; 98US-0090252P.
PR 22-JUN-1998; 98US-0090254P.
PR 24-JUN-1998; 98US-0090429P.
PR 24-JUN-1998; 98US-0090435P.
PR 24-JUN-1998; 98US-0090444P.
PR 24-JUN-1998; 98US-0090461P.
PR 24-JUN-1998; 98US-0090535P.
PR 24-JUN-1998; 98US-0090540P.
PR 25-JUN-1998; 98US-0090676P.
PR 25-JUN-1998; 98US-0090678P.
PR 25-JUN-1998; 98US-0090688P.
PR 25-JUN-1998; 98US-0090690P.
PR 25-JUN-1998; 98US-0090694P.
PR 25-JUN-1998; 98US-0090695P.
PR 25-JUN-1998; 98US-0090696P.
PR 26-JUN-1998; 98US-00105413.
PR 26-JUN-1998; 98US-0090862P.
PR 26-JUN-1998; 98US-0090863P.
PR 26-JUN-1998; 98US-0091010P.
PR 01-JUL-1998; 98US-0091359P.
PR 01-JUL-1998; 98US-0091544P.
PR 02-JUL-1998; 98US-0091478P.
PR 02-JUL-1998; 98US-0091486P.
PR 02-JUL-1998; 98US-0091626P.
PR 02-JUL-1998; 98US-0091628P.
PR 02-JUL-1998; 98US-0091632P.
PR 24-JUL-1998; 98US-0094006P.
PR 04-AUG-1998; 98US-0095282P.
PR 10-AUG-1998; 98US-0095998P.
PR 10-AUG-1998; 98US-0096012P.
PR 17-AUG-1998; 98US-0096757P.
PR 17-AUG-1998; 98US-0096766P.
PR 17-AUG-1998; 98US-0096867P.
PR 17-AUG-1998; 98US-0096891P.
PR 17-AUG-1998; 98US-0096897P.
PR 18-AUG-1998; 98US-0096949P.
PR 18-AUG-1998; 98US-0096959P.
PR 18-AUG-1998; 98US-0097022P.
PR 26-AUG-1998; 98US-0097952P.
PR 26-AUG-1998; 98US-0097954P.
PR 26-AUG-1998; 98US-0097955P.
PR 26-AUG-1998; 98US-0097971P.
PR 26-AUG-1998; 98US-0097974P.
PR 26-AUG-1998; 98US-0098014P.
PR 01-SEP-1998; 98US-0098716P.
PR 01-SEP-1998; 98US-0098723P.
PR 02-SEP-1998; 98US-0098803P.
PR 02-SEP-1998; 98US-0098821P.
PR 02-SEP-1998; 98US-0098843P.
PR 09-SEP-1998; 98US-0099602P.
PR 10-SEP-1998; 98US-0099741P.
PR 10-SEP-1998; 98US-0099754P.
PR 10-SEP-1998; 98US-0099763P.
PR 10-SEP-1998; 98US-0099812P.
PR 15-SEP-1998; 98US-0100388P.
PR 16-SEP-1998; 98US-0100662P.
PR 16-SEP-1998; 98US-0100664P.
PR 16-SEP-1998; 98US-0101751P.
PR 16-SEP-1998; 98WO-US019330.
PR 17-SEP-1998; 98US-0100683P.
PR 17-SEP-1998; 98US-0100684P.
PR 17-SEP-1998; 98US-0100919P.
PR 17-SEP-1998; 98US-0100930P.
PR 18-SEP-1998; 98US-0100849P.
PR 18-SEP-1998; 98US-0101014P.
PR 18-SEP-1998; 98US-0101068P.
PR 23-SEP-1998; 98US-0101471P.
PR 23-SEP-1998; 98US-0101472P.
PR 23-SEP-1998; 98US-0101475P.
PR 23-SEP-1998; 98US-0101477P.
PR 24-SEP-1998; 98US-0101738P.
PR 24-SEP-1998; 98US-0101739P.
PR 24-SEP-1998; 98US-0101743P.
PR 24-SEP-1998; 98US-0101922P.
PR 25-SEP-1998; 98US-0101786P.
PR 29-SEP-1998; 98US-0102207P.
PR 29-SEP-1998; 98US-0102240P.
PR 29-SEP-1998; 98US-0102330P.
PR 29-SEP-1998; 98US-0102331P.
PR 30-SEP-1998; 98US-0102487P.
PR 30-SEP-1998; 98US-0102570P.
PR 30-SEP-1998; 98US-0102571P.
PR 01-OCT-1998; 98US-0102684P.
PR 01-OCT-1998; 98US-0102687P.
PR 02-OCT-1998; 98US-0102965P.
PR 06-OCT-1998; 98US-0103258P.
PR 06-OCT-1998; 98US-0103449P.
PR 07-OCT-1998; 98US-00168978.

Query Match 98.7%; Score 1101; DB 6; Length 382;
Best Local Similarity 96.3%; Pred. No. 5.2e-103;
Matches 206; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

```
Qy          1 ATGRLLS--------GQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 52
              |||||||        ||||||||||||||||||||||||||||||||||||||||||||
Db         22 ATGRLLSASDLDLRGGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 81

Qy         53 ESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVD 112
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 ESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVD 141

Qy        113 EPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGE 172
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 EPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGE 201

Qy        173 ETELTTPVLPEETQEEDAKKTFKESREAALNLAY 206
              ||||||||||||||||||||||||||||||||||
Db        202 ETELTTPVLPEETQEEDAKKTFKESREAALNLAY 235
```

RESULT 22
ABU71468
ID ABU71468 standard; protein; 382 AA.
XX
AC ABU71468;
XX
DT 10-JUN-2003 (first entry)
XX
DE Human PRO polypeptide #24.
XX
KW Human; secreted and transmembrane protein; PRO polypeptide; cancer;
KW Alzheimer's disease; ischaemia; cytostatic; nootropic; vasotropic;
KW neuroprotective.
XX
OS Homo sapiens.
XX
PN US2002192659-A1.
XX
PD 19-DEC-2002.
XX
PF 10-JUL-2001; 2001US-00902853.
XX
PR 17-SEP-1997; 97US-0059113P.
PR 17-SEP-1997; 97US-0059115P.
PR 17-SEP-1997; 97US-0059117P.
PR 17-SEP-1997; 97US-0059119P.
PR 17-SEP-1997; 97US-0059121P.
PR 17-SEP-1997; 97US-0059122P.
PR 17-SEP-1997; 97US-0059184P.
PR 18-SEP-1997; 97US-0059263P.
PR 18-SEP-1997; 97US-0059266P.
PR 15-OCT-1997; 97US-0062125P.
PR 17-OCT-1997; 97US-0062285P.

PR 17-OCT-1997; 97US-0062287P.
PR 21-OCT-1997; 97US-0063486P.
PR 24-OCT-1997; 97US-0062814P.
PR 24-OCT-1997; 97US-0062816P.
PR 24-OCT-1997; 97US-0063045P.
PR 24-OCT-1997; 97US-0063120P.
PR 24-OCT-1997; 97US-0063121P.
PR 24-OCT-1997; 97US-0063127P.
PR 24-OCT-1997; 97US-0063128P.
PR 27-OCT-1997; 97US-0063327P.
PR 27-OCT-1997; 97US-0063329P.
PR 28-OCT-1997; 97US-0063541P.
PR 28-OCT-1997; 97US-0063542P.
PR 28-OCT-1997; 97US-0063544P.
PR 28-OCT-1997; 97US-0063549P.
PR 28-OCT-1997; 97US-0063550P.
PR 28-OCT-1997; 97US-0063564P.
PR 29-OCT-1997; 97US-0063435P.
PR 29-OCT-1997; 97US-0063704P.
PR 29-OCT-1997; 97US-0063732P.
PR 29-OCT-1997; 97US-0063734P.
PR 29-OCT-1997; 97US-0063735P.
PR 29-OCT-1997; 97US-0063738P.
PR 29-OCT-1997; 97US-0064215P.
PR 31-OCT-1997; 97US-0063870P.
PR 31-OCT-1997; 97US-0064103P.
PR 03-NOV-1997; 97US-0064248P.
PR 07-NOV-1997; 97US-0064809P.
PR 12-NOV-1997; 97US-0065186P.
PR 17-NOV-1997; 97US-0065846P.
PR 18-NOV-1997; 97US-0065693P.
PR 21-NOV-1997; 97US-0066120P.
PR 21-NOV-1997; 97US-0066364P.
PR 24-NOV-1997; 97US-0066453P.
PR 24-NOV-1997; 97US-0066466P.
PR 24-NOV-1997; 97US-0066511P.
PR 24-NOV-1997; 97US-0066770P.
PR 24-NOV-1997; 97US-0066772P.
PR 10-SEP-1998; 98WO-US018824.
PR 14-SEP-1998; 98WO-US019177.
PR 16-SEP-1998; 98WO-US019330.
PR 17-SEP-1998; 98WO-US019437.
PR 01-DEC-1998; 98WO-US025108.
PR 08-SEP-1999; 99WO-US020594.
PR 13-SEP-1999; 99WO-US020944.
PR 15-SEP-1999; 99WO-US021090.
PR 15-SEP-1999; 99WO-US021547.
PR 05-OCT-1999; 99WO-US023089.
PR 29-NOV-1999; 99WO-US028214.
PR 30-NOV-1999; 99WO-US028313.
PR 01-DEC-1999; 99WO-US028301.
PR 02-DEC-1999; 99WO-US028564.
PR 02-DEC-1999; 99WO-US028565.
PR 16-DEC-1999; 99WO-US030095.
PR 20-DEC-1999; 99WO-US030911.
PR 20-DEC-1999; 99WO-US030999.
PR 05-JAN-2000; 2000WO-US000219.
PR 11-FEB-2000; 2000WO-US003565.
PR 22-FEB-2000; 2000WO-US004414.
PR 24-FEB-2000; 2000WO-US005004.
PR 02-MAR-2000; 2000WO-US005841.
PR 20-MAR-2000; 2000WO-US007377.
PR 30-MAR-2000; 2000WO-US008439.
PR 22-MAY-2000; 2000WO-US014042.
PR 02-JUN-2000; 2000WO-US015264.
PR 28-JUL-2000; 2000WO-US020710.
PR 24-AUG-2000; 2000WO-US023328.
PR 18-SEP-2000; 2000US-00665350.
XX
PA (GETH) GENENTECH INC.
XX
PI Ashkenazi A, Botstein D, Desnoyers L, Eaton DL, Ferrara N;
PI Filvaroff E, Fong S, Gao W, Gerber H, Gerritsen ME, Goddard A;
PI Godowski PJ, Grimaldi JC, Gurney AL, Hillan KJ, Kljavin IJ;
PI Mather JP, Pan J, Paoni NF, Roy MA, Stewart TA, Tumas D;
PI Williams PM, Wood WI;
XX
DR WPI; 2003-361832/34.
DR N-PSDB; ACA58405.
XX
PT New isolated nucleic acid encoding a PRO polypeptide, e.g. PRO245 or
PT PRO1868, useful in molecular biology, chromosome and gene mapping, in
PT generating antisense RNA and DNA, and in gene therapy.
XX
PS Claim 12; Fig 50; 474pp; English.
XX
CC The present invention relates to the isolation of novel human secreted
CC and transmembrane proteins (PRO polypeptides), and the polynucleotide
CC sequences encoding them. The polynucleotide sequences are useful in
CC molecular biology, as hybridisation probes, in chromosome and gene
CC mapping, in generating antisense RNA and DNA, and in gene therapy. The
CC polynucleotide sequences may also be used in preparing PRO polypeptides
CC by recombinant techniques, and in generating either transgenic animals or
CC knock-out animals which, in turn, are useful in the development and
CC screening of therapeutically useful reagents. The PRO polypeptides or
CC their antibodies are useful in preparing a medicament for treating a
CC condition responsive to the polypeptide or antibody, such as cancer,
CC Alzheimer's disease or ischaemia, and in various diagnostic assays.
CC ABU71445-ABU71505 represent human PRO polypeptides of the invention
XX
SQ Sequence 382 AA;

Query Match 98.7%; Score 1101; DB 6; Length 382;
Best Local Similarity 96.3%; Pred. No. 5.2e-103;
Matches 206; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

```
Qy          1 ATGRLLS--------GQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 52
              |||||||        ||||||||||||||||||||||||||||||||||||||||||||
Db         22 ATGRLLSASDLDLRGGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 81

Qy         53 ESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVD 112
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 ESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVD 141

Qy        113 EPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGE 172
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 EPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGE 201

Qy        173 ETELTTPVLPEETQEEDAKKTFKESREAALNLAY 206
              |||||||||||||||||||||||||||||||||
Db        202 ETELTTPVLPEETQEEDAKKTFKESREAALNLAY 235
```

RESULT 23
ABR68085
ID ABR68085 standard; protein; 382 AA.
XX
AC ABR68085;
XX
DT 11-AUG-2003 (first entry)
XX
DE Human secreted polypeptide PRO234, SEQ ID NO:20.
XX
KW Human; PRO; secreted protein; transmembrane protein;
KW extracellular domain; tumour necrosis factor-alpha; TNF-alpha;
KW chondrocyte; proliferation; differentiation; cartilage disorder;
KW bone disorder; arthritis; sports injury; cancer; tumour; diagnosis;
KW adrenal tumour; lung; colon; breast; prostate; kidney; rectum; cervix;
KW liver; drug screening; transgenic animal; genetic analysis;
KW antiarthritic; vulnerary; gene therapy.
XX
OS Homo sapiens.
XX
PN US2003027264-A1.
XX
PD 06-FEB-2003.

XX
PF 18-JUN-2002; 2002US-00174579.
XX
PR 18-SEP-1997; 97US-0059263P.
PR 18-SEP-1997; 97US-0059266P.
PR 17-OCT-1997; 97US-0062250P.
PR 21-OCT-1997; 97US-0063486P.
PR 24-OCT-1997; 97US-0063120P.
PR 24-OCT-1997; 97US-0063121P.
PR 28-OCT-1997; 97US-0063540P.
PR 28-OCT-1997; 97US-0063541P.
PR 28-OCT-1997; 97US-0063544P.
PR 28-OCT-1997; 97US-0063564P.
PR 29-OCT-1997; 97US-0063734P.
PR 31-OCT-1997; 97US-0063870P.
PR 31-OCT-1997; 97US-0064103P.
PR 13-NOV-1997; 97US-0065311P.
PR 21-NOV-1997; 97US-0066120P.
PR 24-NOV-1997; 97US-0066466P.
PR 24-NOV-1997; 97US-0066772P.
PR 11-DEC-1997; 97US-0069335P.
PR 12-DEC-1997; 97US-0069425P.
PR 17-DEC-1997; 97US-0069870P.
PR 18-DEC-1997; 97US-0068017P.
PR 10-MAR-1998; 98US-0077450P.
PR 11-MAR-1998; 98US-0077632P.
PR 11-MAR-1998; 98US-0077649P.
PR 20-MAR-1998; 98US-0078886P.
PR 20-MAR-1998; 98US-0078939P.
PR 27-MAR-1998; 98US-0079664P.
PR 27-MAR-1998; 98US-0079786P.
PR 31-MAR-1998; 98US-0080107P.
PR 31-MAR-1998; 98US-0080194P.
PR 01-APR-1998; 98US-0080327P.
PR 01-APR-1998; 98US-0080333P.
PR 08-APR-1998; 98US-0081049P.
PR 08-APR-1998; 98US-0081070P.
PR 09-APR-1998; 98US-0081195P.
PR 15-APR-1998; 98US-0081838P.
PR 21-APR-1998; 98US-0082568P.
PR 21-APR-1998; 98US-0082569P.
PR 22-APR-1998; 98US-0082704P.
PR 22-APR-1998; 98US-0082797P.
PR 28-APR-1998; 98US-0083322P.
PR 29-APR-1998; 98US-0083495P.
PR 29-APR-1998; 98US-0083496P.
PR 29-APR-1998; 98US-0083499P.
PR 29-APR-1998; 98US-0083559P.
PR 05-MAY-1998; 98US-0084366P.
PR 06-MAY-1998; 98US-0084414P.
PR 07-MAY-1998; 98US-0084639P.
PR 07-MAY-1998; 98US-0084640P.
PR 07-MAY-1998; 98US-0084643P.
PR 15-MAY-1998; 98US-0085579P.
PR 15-MAY-1998; 98US-0085580P.
PR 15-MAY-1998; 98US-0085582P.
PR 15-MAY-1998; 98US-0085700P.
PR 18-MAY-1998; 98US-0086023P.
PR 22-MAY-1998; 98US-0086392P.
PR 22-MAY-1998; 98US-0086486P.
PR 28-MAY-1998; 98US-0087098P.
PR 28-MAY-1998; 98US-0087208P.
PR 02-JUN-1998; 98US-0087609P.
PR 02-JUN-1998; 98US-0087759P.
PR 03-JUN-1998; 98US-0087827P.
PR 04-JUN-1998; 98US-0088025P.
PR 04-JUN-1998; 98US-0088028P.
PR 04-JUN-1998; 98US-0088029P.
PR 04-JUN-1998; 98US-0088033P.
PR 04-JUN-1998; 98US-0088326P.
PR 05-JUN-1998; 98US-0088167P.
PR 05-JUN-1998; 98US-0088202P.
PR 05-JUN-1998; 98US-0088212P.
PR 05-JUN-1998; 98US-0088217P.
PR 09-JUN-1998; 98US-0088655P.
PR 10-JUN-1998; 98US-0088722P.
PR 10-JUN-1998; 98US-0088738P.
PR 10-JUN-1998; 98US-0088740P.
PR 10-JUN-1998; 98US-0088811P.
PR 10-JUN-1998; 98US-0088824P.
PR 10-JUN-1998; 98US-0088825P.
PR 10-JUN-1998; 98US-0088826P.
PR 11-JUN-1998; 98US-0088861P.
PR 11-JUN-1998; 98US-0088863P.
PR 11-JUN-1998; 98US-0088876P.
PR 12-JUN-1998; 98US-0089090P.
PR 12-JUN-1998; 98US-0089105P.
PR 16-JUN-1998; 98US-0089512P.
PR 16-JUN-1998; 98US-0089514P.
PR 17-JUN-1998; 98US-0089538P.
PR 17-JUN-1998; 98US-0089598P.
PR 17-JUN-1998; 98US-0089653P.
PR 18-JUN-1998; 98US-0089908P.
PR 19-JUN-1998; 98US-0089952P.
PR 22-JUN-1998; 98US-0090246P.
PR 22-JUN-1998; 98US-0090252P.
PR 22-JUN-1998; 98US-0090254P.
PR 24-JUN-1998; 98US-0090429P.
PR 24-JUN-1998; 98US-0090435P.
PR 24-JUN-1998; 98US-0090444P.
PR 24-JUN-1998; 98US-0090461P.
PR 24-JUN-1998; 98US-0090535P.
PR 24-JUN-1998; 98US-0090540P.
PR 25-JUN-1998; 98US-0090676P.
PR 25-JUN-1998; 98US-0090678P.
PR 25-JUN-1998; 98US-0090688P.
PR 25-JUN-1998; 98US-0090690P.
PR 25-JUN-1998; 98US-0090694P.
PR 25-JUN-1998; 98US-0090695P.
PR 25-JUN-1998; 98US-0090696P.
PR 26-JUN-1998; 98US-00105413.
PR 26-JUN-1998; 98US-0090862P.
PR 26-JUN-1998; 98US-0090863P.
PR 26-JUN-1998; 98US-0091010P.
PR 01-JUL-1998; 98US-0091359P.
PR 01-JUL-1998; 98US-0091544P.
PR 02-JUL-1998; 98US-0091478P.
PR 02-JUL-1998; 98US-0091486P.
PR 02-JUL-1998; 98US-0091626P.
PR 02-JUL-1998; 98US-0091628P.
PR 02-JUL-1998; 98US-0091632P.
PR 24-JUL-1998; 98US-0094006P.
PR 04-AUG-1998; 98US-0095282P.
PR 10-AUG-1998; 98US-0095998P.
PR 10-AUG-1998; 98US-0096012P.
PR 17-AUG-1998; 98US-0096757P.
PR 17-AUG-1998; 98US-0096766P.
PR 17-AUG-1998; 98US-0096867P.
PR 17-AUG-1998; 98US-0096891P.
PR 17-AUG-1998; 98US-0096897P.
PR 18-AUG-1998; 98US-0096949P.
PR 18-AUG-1998; 98US-0096959P.
PR 18-AUG-1998; 98US-0097022P.
PR 26-AUG-1998; 98US-0097952P.
PR 26-AUG-1998; 98US-0097954P.
PR 26-AUG-1998; 98US-0097955P.
PR 26-AUG-1998; 98US-0097971P.
PR 26-AUG-1998; 98US-0097974P.
PR 26-AUG-1998; 98US-0098014P.
PR 01-SEP-1998; 98US-0098716P.
PR 01-SEP-1998; 98US-0098723P.
PR 02-SEP-1998; 98US-0098803P.
PR 02-SEP-1998; 98US-0098821P.
PR 02-SEP-1998; 98US-0098843P.
PR 09-SEP-1998; 98US-0099602P.
PR 10-SEP-1998; 98US-0099741P.

PR 10-SEP-1998; 98US-0099754P.
PR 10-SEP-1998; 98US-0099763P.
PR 10-SEP-1998; 98US-0099812P.
PR 15-SEP-1998; 98US-0100388P.
PR 16-SEP-1998; 98US-0100662P.
PR 16-SEP-1998; 98US-0100664P.
PR 16-SEP-1998; 98US-0101751P.
PR 16-SEP-1998; 98WO-US019330.
PR 17-SEP-1998; 98US-0100683P.
PR 17-SEP-1998; 98US-0100684P.
PR 17-SEP-1998; 98US-0100919P.
PR 17-SEP-1998; 98US-0100930P.
PR 18-SEP-1998; 98US-0100849P.
PR 18-SEP-1998; 98US-0101014P.
PR 18-SEP-1998; 98US-0101068P.
PR 23-SEP-1998; 98US-0101471P.
PR 23-SEP-1998; 98US-0101472P.
PR 23-SEP-1998; 98US-0101475P.
PR 23-SEP-1998; 98US-0101477P.
PR 24-SEP-1998; 98US-0101738P.
PR 24-SEP-1998; 98US-0101739P.
PR 24-SEP-1998; 98US-0101743P.
PR 24-SEP-1998; 98US-0101922P.
PR 25-SEP-1998; 98US-0101786P.
PR 29-SEP-1998; 98US-0102207P.
PR 29-SEP-1998; 98US-0102240P.
PR 29-SEP-1998; 98US-0102330P.
PR 29-SEP-1998; 98US-0102331P.
PR 30-SEP-1998; 98US-0102487P.
PR 30-SEP-1998; 98US-0102570P.
PR 30-SEP-1998; 98US-0102571P.
PR 01-OCT-1998; 98US-0102684P.
PR 01-OCT-1998; 98US-0102687P.

Query Match 98.7%; Score 1101; DB 6; Length 382;
Best Local Similarity 96.3%; Pred. No. 5.2e-103;
Matches 206; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

```
Qy          1 ATGRLLS--------GQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 52
              |||||||        ||||||||||||||||||||||||||||||||||||||||||||
Db         22 ATGRLLSASDLDLRGGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 81

Qy         53 ESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVD 112
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 ESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVD 141

Qy        113 EPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGE 172
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 EPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGE 201

Qy        173 ETELTTPVLPEETQEEDAKKTFKESREAALNLAY 206
              |||||||||||||||||||||||||||||||||
Db        202 ETELTTPVLPEETQEEDAKKTFKESREAALNLAY 235
```

RESULT 24
ABU96138
ID ABU96138 standard; protein; 382 AA.
XX
AC ABU96138;
XX
DT 25-JUL-2003 (first entry)
XX
DE Novel human secreted and transmembrane protein PRO234.
XX
KW Human; secreted and transmembrane protein; PRO; transgenic animal;
KW knockout; chromosome identification; tissue typing; tumour;
KW chondrocyte proliferation; chondrocyte differentiation;
KW tumor necrosis factor-alpha release stimulator.
XX
OS Homo sapiens.
XX
PN US2003036144-A1.

XX
PD 20-FEB-2003.
XX
PF 01-JUL-2002; 2002US-00187601.
XX
PR 18-SEP-1997; 97US-0059263P.
PR 18-SEP-1997; 97US-0059266P.
PR 17-OCT-1997; 97US-0062250P.
PR 21-OCT-1997; 97US-0063486P.
PR 24-OCT-1997; 97US-0063120P.
PR 24-OCT-1997; 97US-0063121P.
PR 28-OCT-1997; 97US-0063540P.
PR 28-OCT-1997; 97US-0063541P.
PR 28-OCT-1997; 97US-0063544P.
PR 28-OCT-1997; 97US-0063564P.
PR 29-OCT-1997; 97US-0063734P.
PR 31-OCT-1997; 97US-0063870P.
PR 31-OCT-1997; 97US-0064103P.
PR 13-NOV-1997; 97US-0065311P.
PR 21-NOV-1997; 97US-0066120P.
PR 24-NOV-1997; 97US-0066466P.
PR 24-NOV-1997; 97US-0066772P.
PR 11-DEC-1997; 97US-0069335P.
PR 12-DEC-1997; 97US-0069425P.
PR 17-DEC-1997; 97US-0069870P.
PR 18-DEC-1997; 97US-0068017P.
PR 10-MAR-1998; 98US-0077450P.
PR 11-MAR-1998; 98US-0077632P.
PR 11-MAR-1998; 98US-0077649P.
PR 20-MAR-1998; 98US-0078886P.
PR 20-MAR-1998; 98US-0078939P.
PR 27-MAR-1998; 98US-0079664P.
PR 27-MAR-1998; 98US-0079786P.
PR 31-MAR-1998; 98US-0080107P.
PR 31-MAR-1998; 98US-0080194P.
PR 01-APR-1998; 98US-0080327P.
PR 01-APR-1998; 98US-0080333P.
PR 08-APR-1998; 98US-0081049P.
PR 08-APR-1998; 98US-0081070P.
PR 09-APR-1998; 98US-0081195P.
PR 15-APR-1998; 98US-0081838P.
PR 21-APR-1998; 98US-0082568P.
PR 21-APR-1998; 98US-0082569P.
PR 22-APR-1998; 98US-0082704P.
PR 22-APR-1998; 98US-0082797P.
PR 28-APR-1998; 98US-0083322P.
PR 29-APR-1998; 98US-0083495P.
PR 29-APR-1998; 98US-0083496P.
PR 29-APR-1998; 98US-0083499P.
PR 29-APR-1998; 98US-0083559P.
PR 05-MAY-1998; 98US-0084366P.
PR 06-MAY-1998; 98US-0084414P.
PR 07-MAY-1998; 98US-0084639P.
PR 07-MAY-1998; 98US-0084640P.
PR 07-MAY-1998; 98US-0084643P.
PR 15-MAY-1998; 98US-0085579P.
PR 15-MAY-1998; 98US-0085580P.
PR 15-MAY-1998; 98US-0085582P.
PR 15-MAY-1998; 98US-0085700P.
PR 18-MAY-1998; 98US-0086023P.
PR 22-MAY-1998; 98US-0086392P.
PR 22-MAY-1998; 98US-0086486P.
PR 28-MAY-1998; 98US-0087098P.
PR 28-MAY-1998; 98US-0087208P.
PR 02-JUN-1998; 98US-0087609P.
PR 02-JUN-1998; 98US-0087759P.
PR 03-JUN-1998; 98US-0087827P.
PR 04-JUN-1998; 98US-0088025P.
PR 04-JUN-1998; 98US-0088028P.
PR 04-JUN-1998; 98US-0088029P.
PR 04-JUN-1998; 98US-0088033P.
PR 04-JUN-1998; 98US-0088326P.
PR 05-JUN-1998; 98US-0088167P.

PR 05-JUN-1998; 98US-0088202P.
PR 05-JUN-1998; 98US-0088212P.
PR 05-JUN-1998; 98US-0088217P.
PR 09-JUN-1998; 98US-0088655P.
PR 10-JUN-1998; 98US-0088722P.
PR 10-JUN-1998; 98US-0088738P.
PR 10-JUN-1998; 98US-0088740P.
PR 10-JUN-1998; 98US-0088811P.
PR 10-JUN-1998; 98US-0088824P.
PR 10-JUN-1998; 98US-0088825P.
PR 10-JUN-1998; 98US-0088826P.
PR 11-JUN-1998; 98US-0088861P.
PR 11-JUN-1998; 98US-0088863P.
PR 11-JUN-1998; 98US-0088876P.
PR 12-JUN-1998; 98US-0089090P.
PR 12-JUN-1998; 98US-0089105P.
PR 16-JUN-1998; 98US-0089512P.
PR 16-JUN-1998; 98US-0089514P.
PR 17-JUN-1998; 98US-0089538P.
PR 17-JUN-1998; 98US-0089598P.
PR 17-JUN-1998; 98US-0089653P.
PR 18-JUN-1998; 98US-0089908P.
PR 19-JUN-1998; 98US-0089952P.
PR 22-JUN-1998; 98US-0090246P.
PR 22-JUN-1998; 98US-0090252P.
PR 22-JUN-1998; 98US-0090254P.
PR 24-JUN-1998; 98US-0090429P.
PR 24-JUN-1998; 98US-0090435P.
PR 24-JUN-1998; 98US-0090444P.
PR 24-JUN-1998; 98US-0090461P.
PR 24-JUN-1998; 98US-0090535P.
PR 24-JUN-1998; 98US-0090540P.
PR 25-JUN-1998; 98US-0090676P.
PR 25-JUN-1998; 98US-0090678P.
PR 25-JUN-1998; 98US-0090688P.
PR 25-JUN-1998; 98US-0090690P.
PR 25-JUN-1998; 98US-0090694P.
PR 25-JUN-1998; 98US-0090695P.
PR 25-JUN-1998; 98US-0090696P.
PR 26-JUN-1998; 98US-00105413.
PR 26-JUN-1998; 98US-0090862P.
PR 26-JUN-1998; 98US-0090863P.
PR 26-JUN-1998; 98US-0091010P.
PR 01-JUL-1998; 98US-0091359P.
PR 01-JUL-1998; 98US-0091544P.
PR 02-JUL-1998; 98US-0091478P.
PR 02-JUL-1998; 98US-0091486P.
PR 02-JUL-1998; 98US-0091626P.
PR 02-JUL-1998; 98US-0091628P.
PR 02-JUL-1998; 98US-0091632P.
PR 24-JUL-1998; 98US-0094006P.
PR 04-AUG-1998; 98US-0095282P.
PR 10-AUG-1998; 98US-0095998P.
PR 10-AUG-1998; 98US-0096012P.
PR 17-AUG-1998; 98US-0096757P.
PR 17-AUG-1998; 98US-0096766P.
PR 17-AUG-1998; 98US-0096867P.
PR 17-AUG-1998; 98US-0096891P.
PR 17-AUG-1998; 98US-0096897P.
PR 18-AUG-1998; 98US-0096949P.
PR 18-AUG-1998; 98US-0096959P.
PR 18-AUG-1998; 98US-0097022P.
PR 26-AUG-1998; 98US-0097952P.
PR 26-AUG-1998; 98US-0097954P.
PR 26-AUG-1998; 98US-0097955P.
PR 26-AUG-1998; 98US-0097971P.
PR 26-AUG-1998; 98US-0097974P.
PR 26-AUG-1998; 98US-0098014P.
PR 01-SEP-1998; 98US-0098716P.
PR 01-SEP-1998; 98US-0098723P.
PR 02-SEP-1998; 98US-0098803P.
PR 02-SEP-1998; 98US-0098821P.
PR 02-SEP-1998; 98US-0098843P.
PR 09-SEP-1998; 98US-0099602P.
PR 10-SEP-1998; 98US-0099741P.
PR 10-SEP-1998; 98US-0099754P.
PR 10-SEP-1998; 98US-0099763P.
PR 10-SEP-1998; 98US-0099812P.
PR 15-SEP-1998; 98US-0100388P.
PR 16-SEP-1998; 98US-0100662P.
PR 16-SEP-1998; 98US-0100664P.
PR 16-SEP-1998; 98US-0101751P.
PR 16-SEP-1998; 98WO-US019330.
PR 17-SEP-1998; 98US-0100683P.
PR 17-SEP-1998; 98US-0100684P.
PR 17-SEP-1998; 98US-0100919P.
PR 17-SEP-1998; 98US-0100930P.
PR 18-SEP-1998; 98US-0100849P.
PR 18-SEP-1998; 98US-0101014P.
PR 18-SEP-1998; 98US-0101068P.
PR 23-SEP-1998; 98US-0101471P.
PR 23-SEP-1998; 98US-0101472P.
PR 23-SEP-1998; 98US-0101475P.
PR 23-SEP-1998; 98US-0101477P.
PR 24-SEP-1998; 98US-0101738P.
PR 24-SEP-1998; 98US-0101739P.
PR 24-SEP-1998; 98US-0101743P.
PR 24-SEP-1998; 98US-0101922P.
PR 25-SEP-1998; 98US-0101786P.
PR 29-SEP-1998; 98US-0102207P.
PR 29-SEP-1998; 98US-0102240P.
PR 29-SEP-1998; 98US-0102330P.
PR 29-SEP-1998; 98US-0102331P.
PR 30-SEP-1998; 98US-0102487P.
PR 30-SEP-1998; 98US-0102570P.
PR 30-SEP-1998; 98US-0102571P.
PR 01-OCT-1998; 98US-0102684P.
PR 01-OCT-1998; 98US-0102687P.
PR 02-OCT-1998; 98US-0102965P.
PR 06-OCT-1998; 98US-0103258P.
PR 06-OCT-1998; 98US-0103449P.

Query Match 98.7%; Score 1101; DB 6; Length 382;
Best Local Similarity 96.3%; Pred. No. 5.2e-103;
Matches 206; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

```
Qy          1 ATGRLLS--------GQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 52
              |||||||        ||||||||||||||||||||||||||||||||||||||||||||
Db         22 ATGRLLSASDLDLRGGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 81

Qy         53 ESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVD 112
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 ESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVD 141

Qy        113 EPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGE 172
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 EPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGE 201

Qy        173 ETELTTPVLPEETQEEDAKKTFKESREAALNLAY 206
              |||||||||||||||||||||||||||||||||
Db        202 ETELTTPVLPEETQEEDAKKTFKESREAALNLAY 235
```

RESULT 25
ABU92569
ID ABU92569 standard; protein; 382 AA.
XX
AC ABU92569;
XX
DT 18-JUL-2003 (first entry)
XX
DE Human secreted/transmembrane protein (PRO) #10.
XX
KW Human; secreted protein; transmembrane protein; PRO; tumour;
KW proliferation; differentiation; chondrocyte cell; TNF-alpha;
KW tumour necrosis factor-alpha; gene therapy.

XX
OS Homo sapiens.
XX
PN US2003036149-A1.
XX
PD 20-FEB-2003.
XX
PF 02-JUL-2002; 2002US-00187746.
XX
PR 18-SEP-1997; 97US-0059263P.
PR 18-SEP-1997; 97US-0059266P.
PR 17-OCT-1997; 97US-0062250P.
PR 21-OCT-1997; 97US-0063486P.
PR 24-OCT-1997; 97US-0063120P.
PR 24-OCT-1997; 97US-0063121P.
PR 28-OCT-1997; 97US-0063540P.
PR 28-OCT-1997; 97US-0063541P.
PR 28-OCT-1997; 97US-0063544P.
PR 28-OCT-1997; 97US-0063564P.
PR 29-OCT-1997; 97US-0063734P.
PR 31-OCT-1997; 97US-0063870P.
PR 31-OCT-1997; 97US-0064103P.
PR 13-NOV-1997; 97US-0065311P.
PR 21-NOV-1997; 97US-0066120P.
PR 24-NOV-1997; 97US-0066466P.
PR 24-NOV-1997; 97US-0066772P.
PR 11-DEC-1997; 97US-0069335P.
PR 12-DEC-1997; 97US-0069425P.
PR 17-DEC-1997; 97US-0069870P.
PR 18-DEC-1997; 97US-0068017P.
PR 10-MAR-1998; 98US-0077450P.
PR 11-MAR-1998; 98US-0077632P.
PR 11-MAR-1998; 98US-0077649P.
PR 20-MAR-1998; 98US-0078886P.
PR 20-MAR-1998; 98US-0078939P.
PR 27-MAR-1998; 98US-0079664P.
PR 27-MAR-1998; 98US-0079786P.
PR 31-MAR-1998; 98US-0080107P.
PR 31-MAR-1998; 98US-0080194P.
PR 01-APR-1998; 98US-0080327P.
PR 01-APR-1998; 98US-0080333P.
PR 08-APR-1998; 98US-0081049P.
PR 08-APR-1998; 98US-0081070P.
PR 09-APR-1998; 98US-0081195P.
PR 15-APR-1998; 98US-0081838P.
PR 21-APR-1998; 98US-0082568P.
PR 21-APR-1998; 98US-0082569P.
PR 22-APR-1998; 98US-0082704P.
PR 22-APR-1998; 98US-0082797P.
PR 28-APR-1998; 98US-0083322P.
PR 29-APR-1998; 98US-0083495P.
PR 29-APR-1998; 98US-0083496P.
PR 29-APR-1998; 98US-0083499P.
PR 29-APR-1998; 98US-0083559P.
PR 05-MAY-1998; 98US-0084366P.
PR 06-MAY-1998; 98US-0084414P.
PR 07-MAY-1998; 98US-0084639P.
PR 07-MAY-1998; 98US-0084640P.
PR 07-MAY-1998; 98US-0084643P.
PR 15-MAY-1998; 98US-0085579P.
PR 15-MAY-1998; 98US-0085580P.
PR 15-MAY-1998; 98US-0085582P.
PR 15-MAY-1998; 98US-0085700P.
PR 18-MAY-1998; 98US-0086023P.
PR 22-MAY-1998; 98US-0086392P.
PR 22-MAY-1998; 98US-0086486P.
PR 28-MAY-1998; 98US-0087098P.
PR 28-MAY-1998; 98US-0087208P.
PR 02-JUN-1998; 98US-0087609P.
PR 02-JUN-1998; 98US-0087759P.
PR 03-JUN-1998; 98US-0087827P.
PR 04-JUN-1998; 98US-0088025P.
PR 04-JUN-1998; 98US-0088028P.
PR 04-JUN-1998; 98US-0088029P.
PR 04-JUN-1998; 98US-0088033P.
PR 04-JUN-1998; 98US-0088326P.
PR 05-JUN-1998; 98US-0088167P.
PR 05-JUN-1998; 98US-0088202P.
PR 05-JUN-1998; 98US-0088212P.
PR 05-JUN-1998; 98US-0088217P.
PR 09-JUN-1998; 98US-0088655P.
PR 10-JUN-1998; 98US-0088722P.
PR 10-JUN-1998; 98US-0088738P.
PR 10-JUN-1998; 98US-0088740P.
PR 10-JUN-1998; 98US-0088811P.
PR 10-JUN-1998; 98US-0088824P.
PR 10-JUN-1998; 98US-0088825P.
PR 10-JUN-1998; 98US-0088826P.
PR 11-JUN-1998; 98US-0088861P.
PR 11-JUN-1998; 98US-0088863P.
PR 11-JUN-1998; 98US-0088876P.
PR 12-JUN-1998; 98US-0089090P.
PR 12-JUN-1998; 98US-0089105P.
PR 16-JUN-1998; 98US-0089512P.
PR 16-JUN-1998; 98US-0089514P.
PR 17-JUN-1998; 98US-0089538P.
PR 17-JUN-1998; 98US-0089598P.
PR 17-JUN-1998; 98US-0089653P.
PR 18-JUN-1998; 98US-0089908P.
PR 19-JUN-1998; 98US-0089952P.
PR 22-JUN-1998; 98US-0090246P.
PR 22-JUN-1998; 98US-0090252P.
PR 22-JUN-1998; 98US-0090254P.
PR 24-JUN-1998; 98US-0090429P.
PR 24-JUN-1998; 98US-0090435P.
PR 24-JUN-1998; 98US-0090444P.
PR 24-JUN-1998; 98US-0090461P.
PR 24-JUN-1998; 98US-0090535P.
PR 24-JUN-1998; 98US-0090540P.
PR 25-JUN-1998; 98US-0090676P.
PR 25-JUN-1998; 98US-0090678P.
PR 25-JUN-1998; 98US-0090688P.
PR 25-JUN-1998; 98US-0090690P.
PR 25-JUN-1998; 98US-0090694P.
PR 25-JUN-1998; 98US-0090695P.
PR 25-JUN-1998; 98US-0090696P.
PR 26-JUN-1998; 98US-00105413.
PR 26-JUN-1998; 98US-0090862P.
PR 26-JUN-1998; 98US-0090863P.
PR 26-JUN-1998; 98US-0091010P.
PR 01-JUL-1998; 98US-0091359P.
PR 01-JUL-1998; 98US-0091544P.
PR 02-JUL-1998; 98US-0091478P.
PR 02-JUL-1998; 98US-0091486P.
PR 02-JUL-1998; 98US-0091626P.
PR 02-JUL-1998; 98US-0091628P.
PR 02-JUL-1998; 98US-0091632P.
PR 24-JUL-1998; 98US-0094006P.
PR 04-AUG-1998; 98US-0095282P.
PR 10-AUG-1998; 98US-0095998P.
PR 10-AUG-1998; 98US-0096012P.
PR 17-AUG-1998; 98US-0096757P.
PR 17-AUG-1998; 98US-0096766P.
PR 17-AUG-1998; 98US-0096867P.
PR 17-AUG-1998; 98US-0096891P.
PR 17-AUG-1998; 98US-0096897P.
PR 18-AUG-1998; 98US-0096949P.
PR 18-AUG-1998; 98US-0096959P.
PR 18-AUG-1998; 98US-0097022P.
PR 26-AUG-1998; 98US-0097952P.
PR 26-AUG-1998; 98US-0097954P.
PR 26-AUG-1998; 98US-0097955P.
PR 26-AUG-1998; 98US-0097971P.
PR 26-AUG-1998; 98US-0097974P.
PR 26-AUG-1998; 98US-0098014P.
PR 01-SEP-1998; 98US-0098716P.

PR 01-SEP-1998; 98US-0098723P.
PR 02-SEP-1998; 98US-0098803P.
PR 02-SEP-1998; 98US-0098821P.
PR 02-SEP-1998; 98US-0098843P.
PR 09-SEP-1998; 98US-0099602P.
PR 10-SEP-1998; 98US-0099741P.
PR 10-SEP-1998; 98US-0099754P.
PR 10-SEP-1998; 98US-0099763P.
PR 10-SEP-1998; 98US-0099812P.
PR 15-SEP-1998; 98US-0100388P.
PR 16-SEP-1998; 98US-0100662P.
PR 16-SEP-1998; 98US-0100664P.
PR 16-SEP-1998; 98US-0101751P.
PR 16-SEP-1998; 98WO-US019330.
PR 17-SEP-1998; 98US-0100683P.
PR 17-SEP-1998; 98US-0100684P.
PR 17-SEP-1998; 98US-0100919P.
PR 17-SEP-1998; 98US-0100930P.
PR 18-SEP-1998; 98US-0100849P.
PR 18-SEP-1998; 98US-0101014P.
PR 18-SEP-1998; 98US-0101068P.
PR 23-SEP-1998; 98US-0101471P.
PR 23-SEP-1998; 98US-0101472P.
PR 23-SEP-1998; 98US-0101475P.
PR 23-SEP-1998; 98US-0101477P.
PR 24-SEP-1998; 98US-0101738P.
PR 24-SEP-1998; 98US-0101739P.
PR 24-SEP-1998; 98US-0101743P.
PR 24-SEP-1998; 98US-0101922P.
PR 25-SEP-1998; 98US-0101786P.
PR 29-SEP-1998; 98US-0102207P.
PR 29-SEP-1998; 98US-0102240P.
PR 29-SEP-1998; 98US-0102330P.
PR 29-SEP-1998; 98US-0102331P.
PR 30-SEP-1998; 98US-0102487P.
PR 30-SEP-1998; 98US-0102570P.
PR 30-SEP-1998; 98US-0102571P.
PR 01-OCT-1998; 98US-0102684P.
PR 01-OCT-1998; 98US-0102687P.
PR 02-OCT-1998; 98US-0102965P.
PR 06-OCT-1998; 98US-0103258P.
PR 06-OCT-1998; 98US-0103449P.
PR 07-OCT-1998; 98US-00168978.

Query Match 98.7%; Score 1101; DB 6; Length 382;
Best Local Similarity 96.3%; Pred. No. 5.2e-103;
Matches 206; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

```
Qy          1 ATGRLLS--------GQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 52
              |||||||        ||||||||||||||||||||||||||||||||||||||||||||
Db         22 ATGRLLSASDLDLRGGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 81

Qy         53 ESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVD 112
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 ESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVD 141

Qy        113 EPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGE 172
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 EPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGE 201

Qy        173 ETELTTPVLPEETQEEDAKKTFKESREAALNLAY 206
              |||||||||||||||||||||||||||||||||
Db        202 ETELTTPVLPEETQEEDAKKTFKESREAALNLAY 235
```

RESULT 26
ABO08646
ID ABO08646 standard; protein; 382 AA.
XX
AC ABO08646;
XX
DT 17-AUG-2003 (first entry)
XX
DE Human secreted/transmembrane protein (PRO) #10.
XX
KW Human; secreted and transmembrane protein; PRO; TNF-alpha;
KW tumour necrosis factor alpha; chondrocyte cell; tumour; gene therapy;
KW tissue typing; adrenal tumour; lung tumour; colon tumour; breast tumour;
KW prostate tumour; rectal tumour; cervical tumour; liver tumour.
XX
OS Homo sapiens.
XX
PN US2003044923-A1.
XX
PD 06-MAR-2003.
XX
PF 24-JUN-2002; 2002US-00179522.
XX
PR 18-SEP-1997; 97US-0059263P.
PR 18-SEP-1997; 97US-0059266P.
PR 17-OCT-1997; 97US-0062250P.
PR 21-OCT-1997; 97US-0063486P.
PR 24-OCT-1997; 97US-0063120P.
PR 24-OCT-1997; 97US-0063121P.
PR 28-OCT-1997; 97US-0063540P.
PR 28-OCT-1997; 97US-0063541P.
PR 28-OCT-1997; 97US-0063544P.
PR 28-OCT-1997; 97US-0063564P.
PR 29-OCT-1997; 97US-0063734P.
PR 31-OCT-1997; 97US-0063870P.
PR 31-OCT-1997; 97US-0064103P.
PR 13-NOV-1997; 97US-0065311P.
PR 21-NOV-1997; 97US-0066120P.
PR 24-NOV-1997; 97US-0066466P.
PR 24-NOV-1997; 97US-0066772P.
PR 11-DEC-1997; 97US-0069335P.
PR 12-DEC-1997; 97US-0069425P.
PR 17-DEC-1997; 97US-0069870P.
PR 18-DEC-1997; 97US-0068017P.
PR 10-MAR-1998; 98US-0077450P.
PR 11-MAR-1998; 98US-0077632P.
PR 11-MAR-1998; 98US-0077649P.
PR 20-MAR-1998; 98US-0078886P.
PR 20-MAR-1998; 98US-0078939P.
PR 27-MAR-1998; 98US-0079664P.
PR 27-MAR-1998; 98US-0079786P.
PR 31-MAR-1998; 98US-0080107P.
PR 31-MAR-1998; 98US-0080194P.
PR 01-APR-1998; 98US-0080327P.
PR 01-APR-1998; 98US-0080333P.
PR 08-APR-1998; 98US-0081049P.
PR 08-APR-1998; 98US-0081070P.
PR 09-APR-1998; 98US-0081195P.
PR 15-APR-1998; 98US-0081838P.
PR 21-APR-1998; 98US-0082568P.
PR 21-APR-1998; 98US-0082569P.
PR 22-APR-1998; 98US-0082704P.
PR 22-APR-1998; 98US-0082797P.
PR 28-APR-1998; 98US-0083322P.
PR 29-APR-1998; 98US-0083495P.
PR 29-APR-1998; 98US-0083496P.
PR 29-APR-1998; 98US-0083499P.
PR 29-APR-1998; 98US-0083559P.
PR 05-MAY-1998; 98US-0084366P.
PR 06-MAY-1998; 98US-0084414P.
PR 07-MAY-1998; 98US-0084639P.
PR 07-MAY-1998; 98US-0084640P.
PR 07-MAY-1998; 98US-0084643P.
PR 15-MAY-1998; 98US-0085579P.
PR 15-MAY-1998; 98US-0085580P.
PR 15-MAY-1998; 98US-0085582P.
PR 15-MAY-1998; 98US-0085700P.
PR 18-MAY-1998; 98US-0086023P.
PR 22-MAY-1998; 98US-0086392P.
PR 22-MAY-1998; 98US-0086486P.
PR 28-MAY-1998; 98US-0087098P.

PR 28-MAY-1998; 98US-0087208P.
PR 02-JUN-1998; 98US-0087609P.
PR 02-JUN-1998; 98US-0087759P.
PR 03-JUN-1998; 98US-0087827P.
PR 04-JUN-1998; 98US-0088025P.
PR 04-JUN-1998; 98US-0088028P.
PR 04-JUN-1998; 98US-0088029P.
PR 04-JUN-1998; 98US-0088033P.
PR 04-JUN-1998; 98US-0088326P.
PR 05-JUN-1998; 98US-0088167P.
PR 05-JUN-1998; 98US-0088202P.
PR 05-JUN-1998; 98US-0088212P.
PR 05-JUN-1998; 98US-0088217P.
PR 09-JUN-1998; 98US-0088655P.
PR 10-JUN-1998; 98US-0088722P.
PR 10-JUN-1998; 98US-0088738P.
PR 10-JUN-1998; 98US-0088740P.
PR 10-JUN-1998; 98US-0088811P.
PR 10-JUN-1998; 98US-0088824P.
PR 10-JUN-1998; 98US-0088825P.
PR 10-JUN-1998; 98US-0088826P.
PR 11-JUN-1998; 98US-0088861P.
PR 11-JUN-1998; 98US-0088863P.
PR 11-JUN-1998; 98US-0088876P.
PR 12-JUN-1998; 98US-0089090P.
PR 12-JUN-1998; 98US-0089105P.
PR 16-JUN-1998; 98US-0089512P.
PR 16-JUN-1998; 98US-0089514P.
PR 17-JUN-1998; 98US-0089538P.
PR 17-JUN-1998; 98US-0089598P.
PR 17-JUN-1998; 98US-0089653P.
PR 18-JUN-1998; 98US-0089908P.
PR 19-JUN-1998; 98US-0089952P.
PR 22-JUN-1998; 98US-0090246P.
PR 22-JUN-1998; 98US-0090252P.
PR 22-JUN-1998; 98US-0090254P.
PR 24-JUN-1998; 98US-0090429P.
PR 24-JUN-1998; 98US-0090435P.
PR 24-JUN-1998; 98US-0090444P.
PR 24-JUN-1998; 98US-0090461P.
PR 24-JUN-1998; 98US-0090535P.
PR 24-JUN-1998; 98US-0090540P.
PR 25-JUN-1998; 98US-0090676P.
PR 25-JUN-1998; 98US-0090678P.
PR 25-JUN-1998; 98US-0090688P.
PR 25-JUN-1998; 98US-0090690P.
PR 25-JUN-1998; 98US-0090694P.
PR 25-JUN-1998; 98US-0090695P.
PR 25-JUN-1998; 98US-0090696P.
PR 26-JUN-1998; 98US-00105413.
PR 26-JUN-1998; 98US-0090862P.
PR 26-JUN-1998; 98US-0090863P.
PR 26-JUN-1998; 98US-0091010P.
PR 01-JUL-1998; 98US-0091359P.
PR 01-JUL-1998; 98US-0091544P.
PR 02-JUL-1998; 98US-0091478P.
PR 02-JUL-1998; 98US-0091486P.
PR 02-JUL-1998; 98US-0091626P.
PR 02-JUL-1998; 98US-0091628P.
PR 02-JUL-1998; 98US-0091632P.
PR 24-JUL-1998; 98US-0094006P.
PR 04-AUG-1998; 98US-0095282P.
PR 10-AUG-1998; 98US-0095998P.
PR 10-AUG-1998; 98US-0096012P.
PR 17-AUG-1998; 98US-0096757P.
PR 17-AUG-1998; 98US-0096766P.
PR 17-AUG-1998; 98US-0096867P.
PR 17-AUG-1998; 98US-0096891P.
PR 17-AUG-1998; 98US-0096897P.
PR 18-AUG-1998; 98US-0096949P.
PR 18-AUG-1998; 98US-0096959P.
PR 18-AUG-1998; 98US-0097022P.
PR 26-AUG-1998; 98US-0097952P.
PR 26-AUG-1998; 98US-0097954P.
PR 26-AUG-1998; 98US-0097955P.
PR 26-AUG-1998; 98US-0097971P.
PR 26-AUG-1998; 98US-0097974P.
PR 26-AUG-1998; 98US-0098014P.
PR 01-SEP-1998; 98US-0098716P.
PR 01-SEP-1998; 98US-0098723P.
PR 02-SEP-1998; 98US-0098803P.
PR 02-SEP-1998; 98US-0098821P.
PR 02-SEP-1998; 98US-0098843P.
PR 09-SEP-1998; 98US-0099602P.
PR 10-SEP-1998; 98US-0099741P.
PR 10-SEP-1998; 98US-0099754P.
PR 10-SEP-1998; 98US-0099763P.
PR 10-SEP-1998; 98US-0099812P.
PR 15-SEP-1998; 98US-0100388P.
PR 16-SEP-1998; 98US-0100662P.
PR 16-SEP-1998; 98US-0100664P.
PR 16-SEP-1998; 98US-0101751P.
PR 16-SEP-1998; 98WO-US019330.
PR 17-SEP-1998; 98US-0100683P.
PR 17-SEP-1998; 98US-0100684P.
PR 17-SEP-1998; 98US-0100919P.
PR 17-SEP-1998; 98US-0100930P.
PR 18-SEP-1998; 98US-0100849P.
PR 18-SEP-1998; 98US-0101014P.
PR 18-SEP-1998; 98US-0101068P.
PR 23-SEP-1998; 98US-0101471P.
PR 23-SEP-1998; 98US-0101472P.
PR 23-SEP-1998; 98US-0101475P.
PR 23-SEP-1998; 98US-0101477P.
PR 24-SEP-1998; 98US-0101738P.
PR 24-SEP-1998; 98US-0101739P.
PR 24-SEP-1998; 98US-0101743P.
PR 24-SEP-1998; 98US-0101922P.
PR 25-SEP-1998; 98US-0101786P.
PR 29-SEP-1998; 98US-0102207P.
PR 29-SEP-1998; 98US-0102240P.
PR 29-SEP-1998; 98US-0102330P.
PR 29-SEP-1998; 98US-0102331P.
PR 30-SEP-1998; 98US-0102487P.
PR 30-SEP-1998; 98US-0102570P.
PR 30-SEP-1998; 98US-0102571P.
PR 01-OCT-1998; 98US-0102684P.
PR 01-OCT-1998; 98US-0102687P.
PR 02-OCT-1998; 98US-0102965P.
PR 06-OCT-1998; 98US-0103258P.
PR 06-OCT-1998; 98US-0103449P.

Query Match 98.7%; Score 1101; DB 6; Length 382;
Best Local Similarity 96.3%; Pred. No. 5.2e-103;
Matches 206; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

```
Qy          1 ATGRLLS--------GQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 52
              |||||||        ||||||||||||||||||||||||||||||||||||||||||||
Db         22 ATGRLLSASDLDLRGGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 81

Qy         53 ESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVD 112
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 ESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVD 141

Qy        113 EPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGE 172
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 EPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGE 201

Qy        173 ETELTTPVLPEETQEEDAKKTFKESREAALNLAY 206
              |||||||||||||||||||||||||||||||||
Db        202 ETELTTPVLPEETQEEDAKKTFKESREAALNLAY 235
```

RESULT 27
ABO02698
ID ABO02698 standard; protein; 382 AA.

XX
AC ABO02698;
XX
DT 09-AUG-2003 (first entry)
XX
DE Human secreted/transmembrane protein (PRO) #10.
XX
KW Human; secreted and transmembrane protein; PRO; TNF-alpha;
KW tumour necrosis factor alpha; chondrocyte cell; tumour; gene therapy;
KW tissue typing; adrenal tumour; lung tumour; colon tumour; breast tumour;
KW prostate tumour; rectal tumour; cervical tumour; liver tumour.
XX
OS Homo sapiens.
XX
PN US2003040062-A1.
XX
PD 27-FEB-2003.
XX
PF 25-JUN-2002; 2002US-00180545.
XX
PR 18-SEP-1997; 97US-0059263P.
PR 18-SEP-1997; 97US-0059266P.
PR 17-OCT-1997; 97US-0062250P.
PR 21-OCT-1997; 97US-0063486P.
PR 24-OCT-1997; 97US-0063120P.
PR 24-OCT-1997; 97US-0063121P.
PR 28-OCT-1997; 97US-0063540P.
PR 28-OCT-1997; 97US-0063541P.
PR 28-OCT-1997; 97US-0063544P.
PR 28-OCT-1997; 97US-0063564P.
PR 29-OCT-1997; 97US-0063734P.
PR 31-OCT-1997; 97US-0063870P.
PR 31-OCT-1997; 97US-0064103P.
PR 13-NOV-1997; 97US-0065311P.
PR 21-NOV-1997; 97US-0066120P.
PR 24-NOV-1997; 97US-0066466P.
PR 24-NOV-1997; 97US-0066772P.
PR 11-DEC-1997; 97US-0069335P.
PR 12-DEC-1997; 97US-0069425P.
PR 17-DEC-1997; 97US-0069870P.
PR 18-DEC-1997; 97US-0068017P.
PR 10-MAR-1998; 98US-0077450P.
PR 11-MAR-1998; 98US-0077632P.
PR 11-MAR-1998; 98US-0077649P.
PR 20-MAR-1998; 98US-0078886P.
PR 20-MAR-1998; 98US-0078939P.
PR 27-MAR-1998; 98US-0079664P.
PR 27-MAR-1998; 98US-0079786P.
PR 31-MAR-1998; 98US-0080107P.
PR 31-MAR-1998; 98US-0080194P.
PR 01-APR-1998; 98US-0080327P.
PR 01-APR-1998; 98US-0080333P.
PR 08-APR-1998; 98US-0081049P.
PR 08-APR-1998; 98US-0081070P.
PR 09-APR-1998; 98US-0081195P.
PR 15-APR-1998; 98US-0081838P.
PR 21-APR-1998; 98US-0082568P.
PR 21-APR-1998; 98US-0082569P.
PR 22-APR-1998; 98US-0082704P.
PR 22-APR-1998; 98US-0082797P.
PR 28-APR-1998; 98US-0083322P.
PR 29-APR-1998; 98US-0083495P.
PR 29-APR-1998; 98US-0083496P.
PR 29-APR-1998; 98US-0083499P.
PR 29-APR-1998; 98US-0083559P.
PR 05-MAY-1998; 98US-0084366P.
PR 06-MAY-1998; 98US-0084414P.
PR 07-MAY-1998; 98US-0084639P.
PR 07-MAY-1998; 98US-0084640P.
PR 07-MAY-1998; 98US-0084643P.
PR 15-MAY-1998; 98US-0085579P.
PR 15-MAY-1998; 98US-0085580P.
PR 15-MAY-1998; 98US-0085582P.
PR 15-MAY-1998; 98US-0085700P.
PR 18-MAY-1998; 98US-0086023P.
PR 22-MAY-1998; 98US-0086392P.
PR 22-MAY-1998; 98US-0086486P.
PR 28-MAY-1998; 98US-0087098P.
PR 28-MAY-1998; 98US-0087208P.
PR 02-JUN-1998; 98US-0087609P.
PR 02-JUN-1998; 98US-0087759P.
PR 03-JUN-1998; 98US-0087827P.
PR 04-JUN-1998; 98US-0088025P.
PR 04-JUN-1998; 98US-0088028P.
PR 04-JUN-1998; 98US-0088029P.
PR 04-JUN-1998; 98US-0088033P.
PR 04-JUN-1998; 98US-0088326P.
PR 05-JUN-1998; 98US-0088167P.
PR 05-JUN-1998; 98US-0088202P.
PR 05-JUN-1998; 98US-0088212P.
PR 05-JUN-1998; 98US-0088217P.
PR 09-JUN-1998; 98US-0088655P.
PR 10-JUN-1998; 98US-0088722P.
PR 10-JUN-1998; 98US-0088738P.
PR 10-JUN-1998; 98US-0088740P.
PR 10-JUN-1998; 98US-0088811P.
PR 10-JUN-1998; 98US-0088824P.
PR 10-JUN-1998; 98US-0088825P.
PR 10-JUN-1998; 98US-0088826P.
PR 11-JUN-1998; 98US-0088861P.
PR 11-JUN-1998; 98US-0088863P.
PR 11-JUN-1998; 98US-0088876P.
PR 12-JUN-1998; 98US-0089090P.
PR 12-JUN-1998; 98US-0089105P.
PR 16-JUN-1998; 98US-0089512P.
PR 16-JUN-1998; 98US-0089514P.
PR 17-JUN-1998; 98US-0089538P.
PR 17-JUN-1998; 98US-0089598P.
PR 17-JUN-1998; 98US-0089653P.
PR 18-JUN-1998; 98US-0089908P.
PR 19-JUN-1998; 98US-0089952P.
PR 22-JUN-1998; 98US-0090246P.
PR 22-JUN-1998; 98US-0090252P.
PR 22-JUN-1998; 98US-0090254P.
PR 24-JUN-1998; 98US-0090429P.
PR 24-JUN-1998; 98US-0090435P.
PR 24-JUN-1998; 98US-0090444P.
PR 24-JUN-1998; 98US-0090461P.
PR 24-JUN-1998; 98US-0090535P.
PR 24-JUN-1998; 98US-0090540P.
PR 25-JUN-1998; 98US-0090676P.
PR 25-JUN-1998; 98US-0090678P.
PR 25-JUN-1998; 98US-0090688P.
PR 25-JUN-1998; 98US-0090690P.
PR 25-JUN-1998; 98US-0090694P.
PR 25-JUN-1998; 98US-0090695P.
PR 25-JUN-1998; 98US-0090696P.
PR 26-JUN-1998; 98US-00105413.
PR 26-JUN-1998; 98US-0090862P.
PR 26-JUN-1998; 98US-0090863P.
PR 26-JUN-1998; 98US-0091010P.
PR 01-JUL-1998; 98US-0091359P.
PR 01-JUL-1998; 98US-0091544P.
PR 02-JUL-1998; 98US-0091478P.
PR 02-JUL-1998; 98US-0091486P.
PR 02-JUL-1998; 98US-0091626P.
PR 02-JUL-1998; 98US-0091628P.
PR 02-JUL-1998; 98US-0091632P.
PR 24-JUL-1998; 98US-0094006P.
PR 04-AUG-1998; 98US-0095282P.
PR 10-AUG-1998; 98US-0095998P.
PR 10-AUG-1998; 98US-0096012P.
PR 17-AUG-1998; 98US-0096757P.
PR 17-AUG-1998; 98US-0096766P.
PR 17-AUG-1998; 98US-0096867P.
PR 17-AUG-1998; 98US-0096891P.

PR 17-AUG-1998; 98US-0096897P.
PR 18-AUG-1998; 98US-0096949P.
PR 18-AUG-1998; 98US-0096959P.
PR 18-AUG-1998; 98US-0097022P.
PR 26-AUG-1998; 98US-0097952P.
PR 26-AUG-1998; 98US-0097954P.
PR 26-AUG-1998; 98US-0097955P.
PR 26-AUG-1998; 98US-0097971P.
PR 26-AUG-1998; 98US-0097974P.
PR 26-AUG-1998; 98US-0098014P.
PR 01-SEP-1998; 98US-0098716P.
PR 01-SEP-1998; 98US-0098723P.
PR 02-SEP-1998; 98US-0098803P.
PR 02-SEP-1998; 98US-0098821P.
PR 02-SEP-1998; 98US-0098843P.
PR 09-SEP-1998; 98US-0099602P.
PR 10-SEP-1998; 98US-0099741P.
PR 10-SEP-1998; 98US-0099754P.
PR 10-SEP-1998; 98US-0099763P.
PR 10-SEP-1998; 98US-0099812P.
PR 15-SEP-1998; 98US-0100388P.
PR 16-SEP-1998; 98US-0100662P.
PR 16-SEP-1998; 98US-0100664P.
PR 16-SEP-1998; 98US-0101751P.
PR 16-SEP-1998; 98WO-US019330.
PR 17-SEP-1998; 98US-0100683P.
PR 17-SEP-1998; 98US-0100684P.
PR 17-SEP-1998; 98US-0100919P.
PR 17-SEP-1998; 98US-0100930P.
PR 18-SEP-1998; 98US-0100849P.
PR 18-SEP-1998; 98US-0101014P.
PR 18-SEP-1998; 98US-0101068P.
PR 23-SEP-1998; 98US-0101471P.
PR 23-SEP-1998; 98US-0101472P.
PR 23-SEP-1998; 98US-0101475P.
PR 23-SEP-1998; 98US-0101477P.
PR 24-SEP-1998; 98US-0101738P.
PR 24-SEP-1998; 98US-0101739P.
PR 24-SEP-1998; 98US-0101743P.
PR 24-SEP-1998; 98US-0101922P.
PR 25-SEP-1998; 98US-0101786P.
PR 29-SEP-1998; 98US-0102207P.
PR 29-SEP-1998; 98US-0102240P.
PR 29-SEP-1998; 98US-0102330P.
PR 29-SEP-1998; 98US-0102331P.
PR 30-SEP-1998; 98US-0102487P.
PR 30-SEP-1998; 98US-0102570P.
PR 30-SEP-1998; 98US-0102571P.
PR 01-OCT-1998; 98US-0102684P.
PR 01-OCT-1998; 98US-0102687P.
PR 02-OCT-1998; 98US-0102965P.
PR 06-OCT-1998; 98US-0103258P.
PR 06-OCT-1998; 98US-0103449P.

Query Match 98.7%; Score 1101; DB 6; Length 382;
Best Local Similarity 96.3%; Pred. No. 5.2e-103;
Matches 206; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

```
Qy          1 ATGRLLS--------GQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI  52
              |||||||        ||||||||||||||||||||||||||||||||||||||||||||
Db         22 ATGRLLSASDLDLRGGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI  81

Qy         53 ESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVD  112
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 ESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVD  141

Qy        113 EPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGE  172
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 EPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGE  201

Qy        173 ETELTTPVLPEETQEEDAKKTFKESREAALNLAY  206
              ||||||||||||||||||||||||||||||||
Db        202 ETELTTPVLPEETQEEDAKKTFKESREAALNLAY  235
```

RESULT 28
ABR74852
ID ABR74852 standard; protein; 382 AA.
XX
AC ABR74852;
XX
DT 10-SEP-2003 (first entry)
XX
DE Human secreted polypeptide PRO234, SEQ ID NO:20.
XX
KW Human; PRO; secreted protein; transmembrane protein;
KW extracellular domain; tumour necrosis factor-alpha; TNF-alpha;
KW chondrocyte; proliferation; differentiation; cartilage disorder;
KW bone disorder; arthritis; sports injury; cancer; tumour; diagnosis;
KW adrenal tumour; lung; colon; breast; prostate; kidney; rectum; cervix;
KW liver; drug screening; transgenic animal; genetic analysis;
KW antiarthritic; vulnerary; gene therapy.
XX
OS Homo sapiens.
XX
PN US2003040056-A1.
XX
PD 27-FEB-2003.
XX
PF 21-JUN-2002; 2002US-00176916.
XX
PR 18-SEP-1997; 97US-0059263P.
PR 18-SEP-1997; 97US-0059266P.
PR 17-OCT-1997; 97US-0062250P.
PR 21-OCT-1997; 97US-0063486P.
PR 24-OCT-1997; 97US-0063120P.
PR 24-OCT-1997; 97US-0063121P.
PR 28-OCT-1997; 97US-0063540P.
PR 28-OCT-1997; 97US-0063541P.
PR 28-OCT-1997; 97US-0063544P.
PR 28-OCT-1997; 97US-0063564P.
PR 29-OCT-1997; 97US-0063734P.
PR 31-OCT-1997; 97US-0063870P.
PR 31-OCT-1997; 97US-0064103P.
PR 13-NOV-1997; 97US-0065311P.
PR 21-NOV-1997; 97US-0066120P.
PR 24-NOV-1997; 97US-0066466P.
PR 24-NOV-1997; 97US-0066772P.
PR 11-DEC-1997; 97US-0069335P.
PR 12-DEC-1997; 97US-0069425P.
PR 17-DEC-1997; 97US-0069870P.
PR 18-DEC-1997; 97US-0068017P.
PR 10-MAR-1998; 98US-0077450P.
PR 11-MAR-1998; 98US-0077632P.
PR 11-MAR-1998; 98US-0077649P.
PR 20-MAR-1998; 98US-0078886P.
PR 20-MAR-1998; 98US-0078939P.
PR 27-MAR-1998; 98US-0079664P.
PR 27-MAR-1998; 98US-0079786P.
PR 31-MAR-1998; 98US-0080107P.
PR 31-MAR-1998; 98US-0080194P.
PR 01-APR-1998; 98US-0080327P.
PR 01-APR-1998; 98US-0080333P.
PR 08-APR-1998; 98US-0081049P.
PR 08-APR-1998; 98US-0081070P.
PR 09-APR-1998; 98US-0081195P.
PR 15-APR-1998; 98US-0081838P.
PR 21-APR-1998; 98US-0082568P.
PR 21-APR-1998; 98US-0082569P.
PR 22-APR-1998; 98US-0082704P.
PR 22-APR-1998; 98US-0082797P.
PR 28-APR-1998; 98US-0083322P.
PR 29-APR-1998; 98US-0083495P.
PR 29-APR-1998; 98US-0083496P.
PR 29-APR-1998; 98US-0083499P.
PR 29-APR-1998; 98US-0083559P.

PR 05-MAY-1998; 98US-0084366P.
PR 06-MAY-1998; 98US-0084414P.
PR 07-MAY-1998; 98US-0084639P.
PR 07-MAY-1998; 98US-0084640P.
PR 07-MAY-1998; 98US-0084643P.
PR 15-MAY-1998; 98US-0085579P.
PR 15-MAY-1998; 98US-0085580P.
PR 15-MAY-1998; 98US-0085582P.
PR 15-MAY-1998; 98US-0085700P.
PR 18-MAY-1998; 98US-0086023P.
PR 22-MAY-1998; 98US-0086392P.
PR 22-MAY-1998; 98US-0086486P.
PR 28-MAY-1998; 98US-0087098P.
PR 28-MAY-1998; 98US-0087208P.
PR 02-JUN-1998; 98US-0087609P.
PR 02-JUN-1998; 98US-0087759P.
PR 03-JUN-1998; 98US-0087827P.
PR 04-JUN-1998; 98US-0088025P.
PR 04-JUN-1998; 98US-0088028P.
PR 04-JUN-1998; 98US-0088029P.
PR 04-JUN-1998; 98US-0088033P.
PR 04-JUN-1998; 98US-0088326P.
PR 05-JUN-1998; 98US-0088167P.
PR 05-JUN-1998; 98US-0088202P.
PR 05-JUN-1998; 98US-0088212P.
PR 05-JUN-1998; 98US-0088217P.
PR 09-JUN-1998; 98US-0088655P.
PR 10-JUN-1998; 98US-0088722P.
PR 10-JUN-1998; 98US-0088738P.
PR 10-JUN-1998; 98US-0088740P.
PR 10-JUN-1998; 98US-0088811P.
PR 10-JUN-1998; 98US-0088824P.
PR 10-JUN-1998; 98US-0088825P.
PR 10-JUN-1998; 98US-0088826P.
PR 11-JUN-1998; 98US-0088861P.
PR 11-JUN-1998; 98US-0088863P.
PR 11-JUN-1998; 98US-0088876P.
PR 12-JUN-1998; 98US-0089090P.
PR 12-JUN-1998; 98US-0089105P.
PR 16-JUN-1998; 98US-0089512P.
PR 16-JUN-1998; 98US-0089514P.
PR 17-JUN-1998; 98US-0089538P.
PR 17-JUN-1998; 98US-0089598P.
PR 17-JUN-1998; 98US-0089653P.
PR 18-JUN-1998; 98US-0089908P.
PR 19-JUN-1998; 98US-0089952P.
PR 22-JUN-1998; 98US-0090246P.
PR 22-JUN-1998; 98US-0090252P.
PR 22-JUN-1998; 98US-0090254P.
PR 24-JUN-1998; 98US-0090429P.
PR 24-JUN-1998; 98US-0090435P.
PR 24-JUN-1998; 98US-0090444P.
PR 24-JUN-1998; 98US-0090461P.
PR 24-JUN-1998; 98US-0090535P.
PR 24-JUN-1998; 98US-0090540P.
PR 25-JUN-1998; 98US-0090676P.
PR 25-JUN-1998; 98US-0090678P.
PR 25-JUN-1998; 98US-0090688P.
PR 25-JUN-1998; 98US-0090690P.
PR 25-JUN-1998; 98US-0090694P.
PR 25-JUN-1998; 98US-0090695P.
PR 25-JUN-1998; 98US-0090696P.
PR 26-JUN-1998; 98US-00105413.
PR 26-JUN-1998; 98US-0090862P.
PR 26-JUN-1998; 98US-0090863P.
PR 26-JUN-1998; 98US-0091010P.
PR 01-JUL-1998; 98US-0091359P.
PR 01-JUL-1998; 98US-0091544P.
PR 02-JUL-1998; 98US-0091478P.
PR 02-JUL-1998; 98US-0091486P.
PR 02-JUL-1998; 98US-0091626P.
PR 02-JUL-1998; 98US-0091628P.
PR 02-JUL-1998; 98US-0091632P.
PR 24-JUL-1998; 98US-0094006P.
PR 04-AUG-1998; 98US-0095282P.
PR 10-AUG-1998; 98US-0095998P.
PR 10-AUG-1998; 98US-0096012P.
PR 17-AUG-1998; 98US-0096757P.
PR 17-AUG-1998; 98US-0096766P.
PR 17-AUG-1998; 98US-0096867P.
PR 17-AUG-1998; 98US-0096891P.
PR 17-AUG-1998; 98US-0096897P.
PR 18-AUG-1998; 98US-0096949P.
PR 18-AUG-1998; 98US-0096959P.
PR 18-AUG-1998; 98US-0097022P.
PR 26-AUG-1998; 98US-0097952P.
PR 26-AUG-1998; 98US-0097954P.
PR 26-AUG-1998; 98US-0097955P.
PR 26-AUG-1998; 98US-0097971P.
PR 26-AUG-1998; 98US-0097974P.
PR 26-AUG-1998; 98US-0098014P.
PR 01-SEP-1998; 98US-0098716P.
PR 01-SEP-1998; 98US-0098723P.
PR 02-SEP-1998; 98US-0098803P.
PR 02-SEP-1998; 98US-0098821P.
PR 02-SEP-1998; 98US-0098843P.
PR 09-SEP-1998; 98US-0099602P.
PR 10-SEP-1998; 98US-0099741P.
PR 10-SEP-1998; 98US-0099754P.
PR 10-SEP-1998; 98US-0099763P.
PR 10-SEP-1998; 98US-0099812P.
PR 15-SEP-1998; 98US-0100388P.
PR 16-SEP-1998; 98US-0100662P.
PR 16-SEP-1998; 98US-0100664P.
PR 16-SEP-1998; 98US-0101751P.
PR 16-SEP-1998; 98WO-US019330.
PR 17-SEP-1998; 98US-0100683P.
PR 17-SEP-1998; 98US-0100684P.
PR 17-SEP-1998; 98US-0100919P.
PR 17-SEP-1998; 98US-0100930P.
PR 18-SEP-1998; 98US-0100849P.
PR 18-SEP-1998; 98US-0101014P.
PR 18-SEP-1998; 98US-0101068P.
PR 23-SEP-1998; 98US-0101471P.
PR 23-SEP-1998; 98US-0101472P.
PR 23-SEP-1998; 98US-0101475P.
PR 23-SEP-1998; 98US-0101477P.
PR 24-SEP-1998; 98US-0101738P.
PR 24-SEP-1998; 98US-0101739P.
PR 24-SEP-1998; 98US-0101743P.
PR 24-SEP-1998; 98US-0101922P.
PR 25-SEP-1998; 98US-0101786P.
PR 29-SEP-1998; 98US-0102207P.
PR 29-SEP-1998; 98US-0102240P.
PR 29-SEP-1998; 98US-0102330P.
PR 29-SEP-1998; 98US-0102331P.
PR 30-SEP-1998; 98US-0102487P.
PR 30-SEP-1998; 98US-0102570P.
PR 30-SEP-1998; 98US-0102571P.
PR 01-OCT-1998; 98US-0102684P.
PR 01-OCT-1998; 98US-0102687P.

Query Match 98.7%; Score 1101; DB 6; Length 382;
Best Local Similarity 96.3%; Pred. No. 5.2e-103;
Matches 206; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

```
Qy          1 ATGRLLS--------GQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 52
              |||||||        ||||||||||||||||||||||||||||||||||||||||||||
Db         22 ATGRLLSASDLDLRGGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 81

Qy         53 ESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVD 112
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 ESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVD 141

Qy        113 EPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGE 172
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        142 EPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGE 201

Qy        173 ETELTTPVLPEETQEEDAKKTFKESREAALNLAY 206
              |||||||||||||||||||||||||||||||||
Db        202 ETELTTPVLPEETQEEDAKKTFKESREAALNLAY 235


RESULT 29
ABR94614
ID   ABR94614 standard; protein; 382 AA.
XX
AC   ABR94614;
XX
DT   13-SEP-2003  (first entry)
XX
DE   Human secreted polypeptide PRO234, SEQ ID NO:20.
XX
KW   Human; PRO; secreted protein; transmembrane protein;
KW   extracellular domain; tumour necrosis factor-alpha; TNF-alpha;
KW   chondrocyte; proliferation; differentiation; cartilage disorder;
KW   bone disorder; arthritis; sports injury; cancer; tumour; diagnosis;
KW   adrenal tumour; lung; colon; breast; prostate; kidney; rectum; cervix;
KW   liver; drug screening; transgenic animal; genetic analysis;
KW   antiarthritic; vulnerary; gene therapy.
XX
OS   Homo sapiens.
XX
PN   US2003044926-A1.
XX
PD   06-MAR-2003.
XX
PF   26-JUN-2002; 2002US-00183015.
XX
PR   18-SEP-1997;   97US-0059263P.
PR   18-SEP-1997;   97US-0059266P.
PR   17-OCT-1997;   97US-0062250P.
PR   21-OCT-1997;   97US-0063486P.
PR   24-OCT-1997;   97US-0063120P.
PR   24-OCT-1997;   97US-0063121P.
PR   28-OCT-1997;   97US-0063540P.
PR   28-OCT-1997;   97US-0063541P.
PR   28-OCT-1997;   97US-0063544P.
PR   28-OCT-1997;   97US-0063564P.
PR   29-OCT-1997;   97US-0063734P.
PR   31-OCT-1997;   97US-0063870P.
PR   31-OCT-1997;   97US-0064103P.
PR   13-NOV-1997;   97US-0065311P.
PR   21-NOV-1997;   97US-0066120P.
PR   24-NOV-1997;   97US-0066466P.
PR   24-NOV-1997;   97US-0066772P.
PR   11-DEC-1997;   97US-0069335P.
PR   12-DEC-1997;   97US-0069425P.
PR   17-DEC-1997;   97US-0069870P.
PR   18-DEC-1997;   97US-0068017P.
PR   10-MAR-1998;   98US-0077450P.
PR   11-MAR-1998;   98US-0077632P.
PR   11-MAR-1998;   98US-0077649P.
PR   20-MAR-1998;   98US-0078886P.
PR   20-MAR-1998;   98US-0078939P.
PR   27-MAR-1998;   98US-0079664P.
PR   27-MAR-1998;   98US-0079786P.
PR   31-MAR-1998;   98US-0080107P.
PR   31-MAR-1998;   98US-0080194P.
PR   01-APR-1998;   98US-0080327P.
PR   01-APR-1998;   98US-0080333P.
PR   08-APR-1998;   98US-0081049P.
PR   08-APR-1998;   98US-0081070P.
PR   09-APR-1998;   98US-0081195P.
PR   15-APR-1998;   98US-0081838P.
PR   21-APR-1998;   98US-0082568P.
PR   21-APR-1998;   98US-0082569P.
PR   22-APR-1998;   98US-0082704P.
PR   22-APR-1998;   98US-0082797P.
PR   28-APR-1998;   98US-0083322P.
PR   29-APR-1998;   98US-0083495P.
PR   29-APR-1998;   98US-0083496P.
PR   29-APR-1998;   98US-0083499P.
PR   29-APR-1998;   98US-0083559P.
PR   05-MAY-1998;   98US-0084366P.
PR   06-MAY-1998;   98US-0084414P.
PR   07-MAY-1998;   98US-0084639P.
PR   07-MAY-1998;   98US-0084640P.
PR   07-MAY-1998;   98US-0084643P.
PR   15-MAY-1998;   98US-0085579P.
PR   15-MAY-1998;   98US-0085580P.
PR   15-MAY-1998;   98US-0085582P.
PR   15-MAY-1998;   98US-0085700P.
PR   18-MAY-1998;   98US-0086023P.
PR   22-MAY-1998;   98US-0086392P.
PR   22-MAY-1998;   98US-0086486P.
PR   28-MAY-1998;   98US-0087098P.
PR   28-MAY-1998;   98US-0087208P.
PR   02-JUN-1998;   98US-0087609P.
PR   02-JUN-1998;   98US-0087759P.
PR   03-JUN-1998;   98US-0087827P.
PR   04-JUN-1998;   98US-0088025P.
PR   04-JUN-1998;   98US-0088028P.
PR   04-JUN-1998;   98US-0088029P.
PR   04-JUN-1998;   98US-0088033P.
PR   04-JUN-1998;   98US-0088326P.
PR   05-JUN-1998;   98US-0088167P.
PR   05-JUN-1998;   98US-0088202P.
PR   05-JUN-1998;   98US-0088212P.
PR   05-JUN-1998;   98US-0088217P.
PR   09-JUN-1998;   98US-0088655P.
PR   10-JUN-1998;   98US-0088722P.
PR   10-JUN-1998;   98US-0088738P.
PR   10-JUN-1998;   98US-0088740P.
PR   10-JUN-1998;   98US-0088811P.
PR   10-JUN-1998;   98US-0088824P.
PR   10-JUN-1998;   98US-0088825P.
PR   10-JUN-1998;   98US-0088826P.
PR   11-JUN-1998;   98US-0088861P.
PR   11-JUN-1998;   98US-0088863P.
PR   11-JUN-1998;   98US-0088876P.
PR   12-JUN-1998;   98US-0089090P.
PR   12-JUN-1998;   98US-0089105P.
PR   16-JUN-1998;   98US-0089512P.
PR   16-JUN-1998;   98US-0089514P.
PR   17-JUN-1998;   98US-0089538P.
PR   17-JUN-1998;   98US-0089598P.
PR   17-JUN-1998;   98US-0089653P.
PR   18-JUN-1998;   98US-0089908P.
PR   19-JUN-1998;   98US-0089952P.
PR   22-JUN-1998;   98US-0090246P.
PR   22-JUN-1998;   98US-0090252P.
PR   22-JUN-1998;   98US-0090254P.
PR   24-JUN-1998;   98US-0090429P.
PR   24-JUN-1998;   98US-0090435P.
PR   24-JUN-1998;   98US-0090444P.
PR   24-JUN-1998;   98US-0090461P.
PR   24-JUN-1998;   98US-0090535P.
PR   24-JUN-1998;   98US-0090540P.
PR   25-JUN-1998;   98US-0090676P.
PR   25-JUN-1998;   98US-0090678P.
PR   25-JUN-1998;   98US-0090688P.
PR   25-JUN-1998;   98US-0090690P.
PR   25-JUN-1998;   98US-0090694P.
PR   25-JUN-1998;   98US-0090695P.
PR   25-JUN-1998;   98US-0090696P.
PR   26-JUN-1998;   98US-00105413.
PR   26-JUN-1998;   98US-0090862P.
PR   26-JUN-1998;   98US-0090863P.
PR   26-JUN-1998;   98US-0091010P.
PR   01-JUL-1998;   98US-0091359P.
PR   01-JUL-1998;   98US-0091544P.
```

```
PR   02-JUL-1998;   98US-0091478P.
PR   02-JUL-1998;   98US-0091486P.
PR   02-JUL-1998;   98US-0091626P.
PR   02-JUL-1998;   98US-0091628P.
PR   02-JUL-1998;   98US-0091632P.
PR   24-JUL-1998;   98US-0094006P.
PR   04-AUG-1998;   98US-0095282P.
PR   10-AUG-1998;   98US-0095998P.
PR   10-AUG-1998;   98US-0096012P.
PR   17-AUG-1998;   98US-0096757P.
PR   17-AUG-1998;   98US-0096766P.
PR   17-AUG-1998;   98US-0096867P.
PR   17-AUG-1998;   98US-0096891P.
PR   17-AUG-1998;   98US-0096897P.
PR   18-AUG-1998;   98US-0096949P.
PR   18-AUG-1998;   98US-0096959P.
PR   18-AUG-1998;   98US-0097022P.
PR   26-AUG-1998;   98US-0097952P.
PR   26-AUG-1998;   98US-0097954P.
PR   26-AUG-1998;   98US-0097955P.
PR   26-AUG-1998;   98US-0097971P.
PR   26-AUG-1998;   98US-0097974P.
PR   26-AUG-1998;   98US-0098014P.
PR   01-SEP-1998;   98US-0098716P.
PR   01-SEP-1998;   98US-0098723P.
PR   02-SEP-1998;   98US-0098803P.
PR   02-SEP-1998;   98US-0098821P.
PR   02-SEP-1998;   98US-0098843P.
PR   09-SEP-1998;   98US-0099602P.
PR   10-SEP-1998;   98US-0099741P.
PR   10-SEP-1998;   98US-0099754P.
PR   10-SEP-1998;   98US-0099763P.
PR   10-SEP-1998;   98US-0099812P.
PR   15-SEP-1998;   98US-0100388P.
PR   16-SEP-1998;   98US-0100662P.
PR   16-SEP-1998;   98US-0100664P.
PR   16-SEP-1998;   98US-0101751P.
PR   16-SEP-1998;   98WO-US019330.
PR   17-SEP-1998;   98US-0100683P.
PR   17-SEP-1998;   98US-0100684P.
PR   17-SEP-1998;   98US-0100919P.
PR   17-SEP-1998;   98US-0100930P.
PR   18-SEP-1998;   98US-0100849P.
PR   18-SEP-1998;   98US-0101014P.
PR   18-SEP-1998;   98US-0101068P.
PR   23-SEP-1998;   98US-0101471P.
PR   23-SEP-1998;   98US-0101472P.
PR   23-SEP-1998;   98US-0101475P.
PR   23-SEP-1998;   98US-0101477P.
PR   24-SEP-1998;   98US-0101738P.
PR   24-SEP-1998;   98US-0101739P.
PR   24-SEP-1998;   98US-0101743P.
PR   24-SEP-1998;   98US-0101922P.
PR   25-SEP-1998;   98US-0101786P.
PR   29-SEP-1998;   98US-0102207P.
PR   29-SEP-1998;   98US-0102240P.
PR   29-SEP-1998;   98US-0102330P.
PR   29-SEP-1998;   98US-0102331P.
PR   30-SEP-1998;   98US-0102487P.
PR   30-SEP-1998;   98US-0102570P.
PR   30-SEP-1998;   98US-0102571P.
PR   01-OCT-1998;   98US-0102684P.
PR   01-OCT-1998;   98US-0102687P.

  Query Match             98.7%;  Score 1101;  DB 6;  Length 382;
  Best Local Similarity   96.3%;  Pred. No. 5.2e-103;
  Matches  206;  Conservative    0;  Mismatches    0;  Indels    8;  Gaps    1;

Qy          1 ATGRLLS--------GQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI  52
              |||||||        ||||||||||||||||||||||||||||||||||||||||||||
Db         22 ATGRLLSASDLDLRGGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI  81

Qy         53 ESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVD 112
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 ESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVD 141

Qy        113 EPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGE 172
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 EPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGE 201

Qy        173 ETELTTPVLPEETQEEDAKKTFKESREAALNLAY 206
              |||||||||||||||||||||||||||||||||
Db        202 ETELTTPVLPEETQEEDAKKTFKESREAALNLAY 235


RESULT 30
ABU85587
ID   ABU85587 standard; protein; 382 AA.
XX
AC   ABU85587;
XX
DT   02-JUL-2003  (first entry)
XX
DE   Human PRO polypeptide #10.
XX
KW   Human; PRO; secreted polypeptide; transmembrane polypeptide;
KW   tumour necrosis factor alpha; TNF-alpha; chondrocyte cell; tumour;
KW   cytostatic.
XX
OS   Homo sapiens.
XX
PN   US2003036140-A1.
XX
PD   20-FEB-2003.
XX
PF   01-JUL-2002; 2002US-00187588.
XX
PR   26-JUN-1998;   98US-00105413.
PR   16-SEP-1998;   98WO-US019330.
PR   07-OCT-1998;   98US-00168978.
PR   07-OCT-1998;   98WO-US021141.
PR   06-NOV-1998;   98US-00187368.
PR   01-DEC-1998;   98WO-US025108.
PR   07-DEC-1998;   98US-00202054.
PR   03-MAR-1999;   99US-00254311.
PR   08-MAR-1999;   99WO-US005028.
PR   14-MAY-1999;   99US-00311832.
PR   14-MAY-1999;   99WO-US010733.
PR   02-JUN-1999;   99WO-US012252.
PR   25-AUG-1999;   99US-00380137.
PR   25-AUG-1999;   99US-00380138.
PR   25-AUG-1999;   99US-00380139.
PR   25-AUG-1999;   99US-00380142.
PR   01-SEP-1999;   99WO-US020111.
PR   15-SEP-1999;   99WO-US021090.
PR   18-OCT-1999;   99US-00403297.
PR   12-NOV-1999;   99US-00423844.
PR   01-DEC-1999;   99WO-US028301.
PR   02-DEC-1999;   99WO-US028551.
PR   30-DEC-1999;   99WO-US031274.
PR   05-JAN-2000; 2000WO-US000219.
PR   18-FEB-2000; 2000WO-US004341.
PR   18-FEB-2000; 2000WO-US004342.
PR   22-FEB-2000; 2000WO-US004414.
PR   24-FEB-2000; 2000WO-US005004.
PR   01-MAR-2000; 2000WO-US005601.
PR   02-MAR-2000; 2000WO-US005841.
PR   15-MAR-2000; 2000WO-US006884.
PR   30-MAR-2000; 2000WO-US008439.
PR   17-MAY-2000; 2000WO-US013705.
PR   22-MAY-2000; 2000WO-US014042.
PR   30-MAY-2000; 2000WO-US014941.
PR   02-JUN-2000; 2000WO-US015264.
PR   28-JUL-2000; 2000WO-US020710.
PR   22-AUG-2000; 2000US-00644848.
PR   24-AUG-2000; 2000WO-US023328.
```

PR 18-SEP-2000; 2000US-00664610.
PR 18-SEP-2000; 2000US-00665350.
PR 08-NOV-2000; 2000US-00709238.
PR 08-NOV-2000; 2000WO-US030952.
PR 01-DEC-2000; 2000WO-US032678.
PR 20-DEC-2000; 2000US-00747259.
PR 20-DEC-2000; 2000WO-US034956.
PR 28-FEB-2001; 2001WO-US006520.
PR 22-MAR-2001; 2001US-00816744.
PR 10-MAY-2001; 2001US-00854208.
PR 10-MAY-2001; 2001US-00854280.
PR 25-MAY-2001; 2001US-00866028.
PR 01-JUN-2001; 2001WO-US017800.
PR 05-JUN-2001; 2001US-00874503.
PR 20-JUN-2001; 2001WO-US019692.
PR 29-JUN-2001; 2001WO-US021066.
PR 09-JUL-2001; 2001WO-US021735.
PR 18-JUL-2001; 2001US-00908827.
PR 30-JUL-2001; 2001US-00918585.
PR 06-AUG-2001; 2001US-00924419.
PR 13-AUG-2001; 2001US-00929404.
PR 16-AUG-2001; 2001US-00931836.
PR 28-AUG-2001; 2001US-00941992.
PR 29-AUG-2001; 2001WO-US027099.
PR 04-SEP-2001; 2001US-00946374.
PR 15-JAN-2002; 2002US-00052586.
XX
PA (GETH) GENENTECH INC.
XX
PI Baker KP, Chen J, Desnoyers L, Goddard A, Godowski PJ, Gurney AL;
PI Pan J, Smith V, Watanabe CK, Wood WI, Zhang Z;
XX
DR WPI; 2003-332028/31.
DR N-PSDB; ACA72780.
XX
PT Three hundred and five nucleic acids encoding PRO polypeptides, useful
PT for the manufacture of a medicament for diagnosing or treating tumor.
XX
PS Claim 11; Fig 20; 707pp; English.
XX
CC The invention relates to human PRO polypeptides (secreted and
CC transmembrane polypeptides) and the PRO polynucleotides encoding them.
CC The invention also relates to a method for stimulating the release of
CC tumour necrosis factor alpha (TNF-alpha) from human blood by contacting
CC the blood with a sequence of the invention, a method for stimulating the
CC proliferation or differentiation of chondrocyte cells by contacting the
CC cells with a PRO polypeptide and a method for detecting the presence of a
CC tumour in a mammal. The polypeptides and polynucleotides are useful for
CC the manufacture of a medicament for diagnosing or treating a tumour in a
CC mammal. Sequences ABU85578-ABU85882 represent human PRO polypeptides of
CC the invention. Note: The sequence data for this patent is also available
CC in electronic format from USPTO at seqdata.uspto.gov/sequence.html
XX
SQ Sequence 382 AA;

Query Match 98.7%; Score 1101; DB 6; Length 382;
Best Local Similarity 96.3%; Pred. No. 5.2e-103;
Matches 206; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

```
Qy          1 ATGRLLS--------GQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 52
              |||||||        ||||||||||||||||||||||||||||||||||||||||||||
Db         22 ATGRLLSASDLDLRGGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 81

Qy         53 ESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVD 112
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 ESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVD 141

Qy        113 EPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGE 172
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 EPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGE 201

Qy        173 ETELTTPVLPEETQEEDAKKTFKESREAALNLAY 206
              |||||||||||||||||||||||||||||||||
```

Db 202 ETELTTPVLPEETQEEDAKKTFKESREAALNLAY 235

RESULT 31
ABU98747
ID ABU98747 standard; protein; 382 AA.
XX
AC ABU98747;
XX
DT 01-AUG-2003 (first entry)
XX
DE Novel human secreted and transmembrane protein PRO234.
XX
KW Human; secreted and transmembrane protein; PRO; cytostatic; gene therapy;
KW chondrocyte stimulator; tumour; adrenal tumour; lung tumour;
KW colon tumour; breast tumour; prostate tumour; rectal tumour;
KW cervical tumour; liver tumour; TNF-alpha release;
KW tumour necrosis factor alpha release; chondrocyte cell proliferation;
KW chondrocyte cell differentiation; pharmaceutical; diagnostic; biosensor;
KW bioreactor.
XX
OS Homo sapiens.
XX
PN US2003013153-A1.
XX
PD 16-JAN-2003.
XX
PF 19-JUN-2002; 2002US-00175737.
XX
PR 18-SEP-1997; 97US-0059263P.
PR 18-SEP-1997; 97US-0059266P.
PR 17-OCT-1997; 97US-0062250P.
PR 21-OCT-1997; 97US-0063486P.
PR 24-OCT-1997; 97US-0063120P.
PR 24-OCT-1997; 97US-0063121P.
PR 28-OCT-1997; 97US-0063540P.
PR 28-OCT-1997; 97US-0063541P.
PR 28-OCT-1997; 97US-0063544P.
PR 28-OCT-1997; 97US-0063564P.
PR 29-OCT-1997; 97US-0063734P.
PR 31-OCT-1997; 97US-0063870P.
PR 31-OCT-1997; 97US-0064103P.
PR 13-NOV-1997; 97US-0065311P.
PR 21-NOV-1997; 97US-0066120P.
PR 24-NOV-1997; 97US-0066466P.
PR 24-NOV-1997; 97US-0066772P.
PR 11-DEC-1997; 97US-0069335P.
PR 12-DEC-1997; 97US-0069425P.
PR 17-DEC-1997; 97US-0069870P.
PR 18-DEC-1997; 97US-0068017P.
PR 10-MAR-1998; 98US-0077450P.
PR 11-MAR-1998; 98US-0077632P.
PR 11-MAR-1998; 98US-0077649P.
PR 20-MAR-1998; 98US-0078886P.
PR 20-MAR-1998; 98US-0078939P.
PR 27-MAR-1998; 98US-0079664P.
PR 27-MAR-1998; 98US-0079786P.
PR 31-MAR-1998; 98US-0080107P.
PR 31-MAR-1998; 98US-0080194P.
PR 01-APR-1998; 98US-0080327P.
PR 01-APR-1998; 98US-0080333P.
PR 08-APR-1998; 98US-0081049P.
PR 08-APR-1998; 98US-0081070P.
PR 09-APR-1998; 98US-0081195P.
PR 15-APR-1998; 98US-0081838P.
PR 21-APR-1998; 98US-0082568P.
PR 21-APR-1998; 98US-0082569P.
PR 22-APR-1998; 98US-0082704P.
PR 22-APR-1998; 98US-0082797P.
PR 28-APR-1998; 98US-0083322P.
PR 29-APR-1998; 98US-0083495P.
PR 29-APR-1998; 98US-0083496P.
PR 29-APR-1998; 98US-0083499P.

PR 29-APR-1998; 98US-0083559P.
PR 05-MAY-1998; 98US-0084366P.
PR 06-MAY-1998; 98US-0084414P.
PR 07-MAY-1998; 98US-0084639P.
PR 07-MAY-1998; 98US-0084640P.
PR 07-MAY-1998; 98US-0084643P.
PR 15-MAY-1998; 98US-0085579P.
PR 15-MAY-1998; 98US-0085580P.
PR 15-MAY-1998; 98US-0085582P.
PR 15-MAY-1998; 98US-0085700P.
PR 18-MAY-1998; 98US-0086023P.
PR 22-MAY-1998; 98US-0086392P.
PR 22-MAY-1998; 98US-0086486P.
PR 28-MAY-1998; 98US-0087098P.
PR 28-MAY-1998; 98US-0087208P.
PR 02-JUN-1998; 98US-0087609P.
PR 02-JUN-1998; 98US-0087759P.
PR 03-JUN-1998; 98US-0087827P.
PR 04-JUN-1998; 98US-0088025P.
PR 04-JUN-1998; 98US-0088028P.
PR 04-JUN-1998; 98US-0088029P.
PR 04-JUN-1998; 98US-0088033P.
PR 04-JUN-1998; 98US-0088326P.
PR 05-JUN-1998; 98US-0088167P.
PR 05-JUN-1998; 98US-0088202P.
PR 05-JUN-1998; 98US-0088212P.
PR 05-JUN-1998; 98US-0088217P.
PR 09-JUN-1998; 98US-0088655P.
PR 10-JUN-1998; 98US-0088722P.
PR 10-JUN-1998; 98US-0088738P.
PR 10-JUN-1998; 98US-0088740P.
PR 10-JUN-1998; 98US-0088811P.
PR 10-JUN-1998; 98US-0088824P.
PR 10-JUN-1998; 98US-0088825P.
PR 10-JUN-1998; 98US-0088826P.
PR 11-JUN-1998; 98US-0088861P.
PR 11-JUN-1998; 98US-0088863P.
PR 11-JUN-1998; 98US-0088876P.
PR 12-JUN-1998; 98US-0089090P.
PR 12-JUN-1998; 98US-0089105P.
PR 16-JUN-1998; 98US-0089512P.
PR 16-JUN-1998; 98US-0089514P.
PR 17-JUN-1998; 98US-0089538P.
PR 17-JUN-1998; 98US-0089598P.
PR 17-JUN-1998; 98US-0089653P.
PR 18-JUN-1998; 98US-0089908P.
PR 19-JUN-1998; 98US-0089952P.
PR 22-JUN-1998; 98US-0090246P.
PR 22-JUN-1998; 98US-0090252P.
PR 22-JUN-1998; 98US-0090254P.
PR 24-JUN-1998; 98US-0090429P.
PR 24-JUN-1998; 98US-0090435P.
PR 24-JUN-1998; 98US-0090444P.
PR 24-JUN-1998; 98US-0090461P.
PR 24-JUN-1998; 98US-0090535P.
PR 24-JUN-1998; 98US-0090540P.
PR 25-JUN-1998; 98US-0090676P.
PR 25-JUN-1998; 98US-0090678P.
PR 25-JUN-1998; 98US-0090688P.
PR 25-JUN-1998; 98US-0090690P.
PR 25-JUN-1998; 98US-0090694P.
PR 25-JUN-1998; 98US-0090695P.
PR 25-JUN-1998; 98US-0090696P.
PR 26-JUN-1998; 98US-00105413.
PR 26-JUN-1998; 98US-0090862P.
PR 26-JUN-1998; 98US-0090863P.
PR 26-JUN-1998; 98US-0091010P.
PR 01-JUL-1998; 98US-0091359P.
PR 01-JUL-1998; 98US-0091544P.
PR 02-JUL-1998; 98US-0091478P.
PR 02-JUL-1998; 98US-0091486P.
PR 02-JUL-1998; 98US-0091626P.
PR 02-JUL-1998; 98US-0091628P.
PR 02-JUL-1998; 98US-0091632P.
PR 24-JUL-1998; 98US-0094006P.
PR 04-AUG-1998; 98US-0095282P.
PR 10-AUG-1998; 98US-0095998P.
PR 10-AUG-1998; 98US-0096012P.
PR 17-AUG-1998; 98US-0096757P.
PR 17-AUG-1998; 98US-0096766P.
PR 17-AUG-1998; 98US-0096867P.
PR 17-AUG-1998; 98US-0096891P.
PR 17-AUG-1998; 98US-0096897P.
PR 18-AUG-1998; 98US-0096949P.
PR 18-AUG-1998; 98US-0096959P.
PR 18-AUG-1998; 98US-0097022P.
PR 26-AUG-1998; 98US-0097952P.
PR 26-AUG-1998; 98US-0097954P.
PR 26-AUG-1998; 98US-0097955P.
PR 26-AUG-1998; 98US-0097971P.
PR 26-AUG-1998; 98US-0097974P.
PR 26-AUG-1998; 98US-0098014P.
PR 01-SEP-1998; 98US-0098716P.
PR 01-SEP-1998; 98US-0098723P.
PR 02-SEP-1998; 98US-0098803P.
PR 02-SEP-1998; 98US-0098821P.
PR 02-SEP-1998; 98US-0098843P.
PR 09-SEP-1998; 98US-0099602P.
PR 10-SEP-1998; 98US-0099741P.
PR 10-SEP-1998; 98US-0099754P.
PR 10-SEP-1998; 98US-0099763P.
PR 10-SEP-1998; 98US-0099812P.
PR 15-SEP-1998; 98US-0100388P.
PR 16-SEP-1998; 98US-0100662P.
PR 16-SEP-1998; 98US-0100664P.
PR 16-SEP-1998; 98US-0101751P.
PR 16-SEP-1998; 98WO-US019330.
PR 17-SEP-1998; 98US-0100683P.
PR 17-SEP-1998; 98US-0100684P.
PR 17-SEP-1998; 98US-0100919P.
PR 17-SEP-1998; 98US-0100930P.
PR 18-SEP-1998; 98US-0100849P.
PR 18-SEP-1998; 98US-0101014P.
PR 18-SEP-1998; 98US-0101068P.
PR 23-SEP-1998; 98US-0101471P.
PR 23-SEP-1998; 98US-0101472P.
PR 23-SEP-1998; 98US-0101475P.
PR 23-SEP-1998; 98US-0101477P.
PR 24-SEP-1998; 98US-0101738P.
PR 24-SEP-1998; 98US-0101739P.
PR 24-SEP-1998; 98US-0101743P.
PR 24-SEP-1998; 98US-0101922P.
PR 25-SEP-1998; 98US-0101786P.
PR 29-SEP-1998; 98US-0102207P.
PR 29-SEP-1998; 98US-0102240P.
PR 29-SEP-1998; 98US-0102330P.
PR 29-SEP-1998; 98US-0102331P.
PR 30-SEP-1998; 98US-0102487P.
PR 30-SEP-1998; 98US-0102570P.
PR 30-SEP-1998; 98US-0102571P.
PR 01-OCT-1998; 98US-0102684P.
PR 01-OCT-1998; 98US-0102687P.

Query Match 98.7%; Score 1101; DB 6; Length 382;
Best Local Similarity 96.3%; Pred. No. 5.2e-103;
Matches 206; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

```
Qy          1 ATGRLLS--------GQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 52
              |||||||        ||||||||||||||||||||||||||||||||||||||||||||
Db         22 ATGRLLSASDLDLRGGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 81

Qy         53 ESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVD 112
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 ESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVD 141

Qy        113 EPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGE 172
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 EPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGE 201

Qy        173 ETELTTPVLPEETQEEDAKKTFKESREAALNLAY 206
                |||||||||||||||||||||||||||||||||
Db        202 ETELTTPVLPEETQEEDAKKTFKESREAALNLAY 235


RESULT 32
ABU97962
ID   ABU97962 standard; protein; 382 AA.
XX
AC   ABU97962;
XX
DT   30-JUL-2003  (first entry)
XX
DE   Novel human secreted and transmembrane protein PRO234.
XX
KW   Human; secreted and transmembrane protein; PRO; cytostatic; gene therapy;
KW   chondrocyte stimulator; tumour; adrenal tumour; lung tumour;
KW   colon tumour; breast tumour; prostate tumour; rectal tumour;
KW   cervical tumour; liver tumour; chromosome identification.
XX
OS   Homo sapiens.
XX
PN   US2003017544-A1.
XX
PD   23-JAN-2003.
XX
PF   21-JUN-2002; 2002US-00176915.
XX
PR   18-SEP-1997;   97US-0059263P.
PR   18-SEP-1997;   97US-0059266P.
PR   17-OCT-1997;   97US-0062250P.
PR   21-OCT-1997;   97US-0063486P.
PR   24-OCT-1997;   97US-0063120P.
PR   24-OCT-1997;   97US-0063121P.
PR   28-OCT-1997;   97US-0063540P.
PR   28-OCT-1997;   97US-0063541P.
PR   28-OCT-1997;   97US-0063544P.
PR   28-OCT-1997;   97US-0063564P.
PR   29-OCT-1997;   97US-0063734P.
PR   31-OCT-1997;   97US-0063870P.
PR   31-OCT-1997;   97US-0064103P.
PR   13-NOV-1997;   97US-0065311P.
PR   21-NOV-1997;   97US-0066120P.
PR   24-NOV-1997;   97US-0066466P.
PR   24-NOV-1997;   97US-0066772P.
PR   11-DEC-1997;   97US-0069335P.
PR   12-DEC-1997;   97US-0069425P.
PR   17-DEC-1997;   97US-0069870P.
PR   18-DEC-1997;   97US-0068017P.
PR   10-MAR-1998;   98US-0077450P.
PR   11-MAR-1998;   98US-0077632P.
PR   11-MAR-1998;   98US-0077649P.
PR   20-MAR-1998;   98US-0078886P.
PR   20-MAR-1998;   98US-0078939P.
PR   27-MAR-1998;   98US-0079664P.
PR   27-MAR-1998;   98US-0079786P.
PR   31-MAR-1998;   98US-0080107P.
PR   31-MAR-1998;   98US-0080194P.
PR   01-APR-1998;   98US-0080327P.
PR   01-APR-1998;   98US-0080333P.
PR   08-APR-1998;   98US-0081049P.
PR   08-APR-1998;   98US-0081070P.
PR   09-APR-1998;   98US-0081195P.
PR   15-APR-1998;   98US-0081838P.
PR   21-APR-1998;   98US-0082568P.
PR   21-APR-1998;   98US-0082569P.
PR   22-APR-1998;   98US-0082704P.
PR   22-APR-1998;   98US-0082797P.
PR   28-APR-1998;   98US-0083322P.
PR   29-APR-1998;   98US-0083495P.
PR   29-APR-1998;   98US-0083496P.
PR   29-APR-1998;   98US-0083499P.
PR   29-APR-1998;   98US-0083559P.
PR   05-MAY-1998;   98US-0084366P.
PR   06-MAY-1998;   98US-0084414P.
PR   07-MAY-1998;   98US-0084639P.
PR   07-MAY-1998;   98US-0084640P.
PR   07-MAY-1998;   98US-0084643P.
PR   15-MAY-1998;   98US-0085579P.
PR   15-MAY-1998;   98US-0085580P.
PR   15-MAY-1998;   98US-0085582P.
PR   15-MAY-1998;   98US-0085700P.
PR   18-MAY-1998;   98US-0086023P.
PR   22-MAY-1998;   98US-0086392P.
PR   22-MAY-1998;   98US-0086486P.
PR   28-MAY-1998;   98US-0087098P.
PR   28-MAY-1998;   98US-0087208P.
PR   02-JUN-1998;   98US-0087609P.
PR   02-JUN-1998;   98US-0087759P.
PR   03-JUN-1998;   98US-0087827P.
PR   04-JUN-1998;   98US-0088025P.
PR   04-JUN-1998;   98US-0088028P.
PR   04-JUN-1998;   98US-0088029P.
PR   04-JUN-1998;   98US-0088033P.
PR   04-JUN-1998;   98US-0088326P.
PR   05-JUN-1998;   98US-0088167P.
PR   05-JUN-1998;   98US-0088202P.
PR   05-JUN-1998;   98US-0088212P.
PR   05-JUN-1998;   98US-0088217P.
PR   09-JUN-1998;   98US-0088655P.
PR   10-JUN-1998;   98US-0088722P.
PR   10-JUN-1998;   98US-0088738P.
PR   10-JUN-1998;   98US-0088740P.
PR   10-JUN-1998;   98US-0088811P.
PR   10-JUN-1998;   98US-0088824P.
PR   10-JUN-1998;   98US-0088825P.
PR   10-JUN-1998;   98US-0088826P.
PR   11-JUN-1998;   98US-0088861P.
PR   11-JUN-1998;   98US-0088863P.
PR   11-JUN-1998;   98US-0088876P.
PR   12-JUN-1998;   98US-0089090P.
PR   12-JUN-1998;   98US-0089105P.
PR   16-JUN-1998;   98US-0089512P.
PR   16-JUN-1998;   98US-0089514P.
PR   17-JUN-1998;   98US-0089538P.
PR   17-JUN-1998;   98US-0089598P.
PR   17-JUN-1998;   98US-0089653P.
PR   18-JUN-1998;   98US-0089908P.
PR   19-JUN-1998;   98US-0089952P.
PR   22-JUN-1998;   98US-0090246P.
PR   22-JUN-1998;   98US-0090252P.
PR   22-JUN-1998;   98US-0090254P.
PR   24-JUN-1998;   98US-0090429P.
PR   24-JUN-1998;   98US-0090435P.
PR   24-JUN-1998;   98US-0090444P.
PR   24-JUN-1998;   98US-0090461P.
PR   24-JUN-1998;   98US-0090535P.
PR   24-JUN-1998;   98US-0090540P.
PR   25-JUN-1998;   98US-0090676P.
PR   25-JUN-1998;   98US-0090678P.
PR   25-JUN-1998;   98US-0090688P.
PR   25-JUN-1998;   98US-0090690P.
PR   25-JUN-1998;   98US-0090694P.
PR   25-JUN-1998;   98US-0090695P.
PR   25-JUN-1998;   98US-0090696P.
PR   26-JUN-1998;   98US-00105413.
PR   26-JUN-1998;   98US-0090862P.
PR   26-JUN-1998;   98US-0090863P.
PR   26-JUN-1998;   98US-0091010P.
PR   01-JUL-1998;   98US-0091359P.
PR   01-JUL-1998;   98US-0091544P.
PR   02-JUL-1998;   98US-0091478P.
PR   02-JUL-1998;   98US-0091486P.
```

PR 02-JUL-1998; 98US-0091626P.
PR 02-JUL-1998; 98US-0091628P.
PR 02-JUL-1998; 98US-0091632P.
PR 24-JUL-1998; 98US-0094006P.
PR 04-AUG-1998; 98US-0095282P.
PR 10-AUG-1998; 98US-0095998P.
PR 10-AUG-1998; 98US-0096012P.
PR 17-AUG-1998; 98US-0096757P.
PR 17-AUG-1998; 98US-0096766P.
PR 17-AUG-1998; 98US-0096867P.
PR 17-AUG-1998; 98US-0096891P.
PR 17-AUG-1998; 98US-0096897P.
PR 18-AUG-1998; 98US-0096949P.
PR 18-AUG-1998; 98US-0096959P.
PR 18-AUG-1998; 98US-0097022P.
PR 26-AUG-1998; 98US-0097952P.
PR 26-AUG-1998; 98US-0097954P.
PR 26-AUG-1998; 98US-0097955P.
PR 26-AUG-1998; 98US-0097971P.
PR 26-AUG-1998; 98US-0097974P.
PR 26-AUG-1998; 98US-0098014P.
PR 01-SEP-1998; 98US-0098716P.
PR 01-SEP-1998; 98US-0098723P.
PR 02-SEP-1998; 98US-0098803P.
PR 02-SEP-1998; 98US-0098821P.
PR 02-SEP-1998; 98US-0098843P.
PR 09-SEP-1998; 98US-0099602P.
PR 10-SEP-1998; 98US-0099741P.
PR 10-SEP-1998; 98US-0099754P.
PR 10-SEP-1998; 98US-0099763P.
PR 10-SEP-1998; 98US-0099812P.
PR 15-SEP-1998; 98US-0100388P.
PR 16-SEP-1998; 98US-0100662P.
PR 16-SEP-1998; 98US-0100664P.
PR 16-SEP-1998; 98US-0101751P.
PR 16-SEP-1998; 98WO-US019330.
PR 17-SEP-1998; 98US-0100683P.
PR 17-SEP-1998; 98US-0100684P.
PR 17-SEP-1998; 98US-0100919P.
PR 17-SEP-1998; 98US-0100930P.
PR 18-SEP-1998; 98US-0100849P.
PR 18-SEP-1998; 98US-0101014P.
PR 18-SEP-1998; 98US-0101068P.
PR 23-SEP-1998; 98US-0101471P.
PR 23-SEP-1998; 98US-0101472P.
PR 23-SEP-1998; 98US-0101475P.
PR 23-SEP-1998; 98US-0101477P.
PR 24-SEP-1998; 98US-0101738P.
PR 24-SEP-1998; 98US-0101739P.
PR 24-SEP-1998; 98US-0101743P.
PR 24-SEP-1998; 98US-0101922P.
PR 25-SEP-1998; 98US-0101786P.
PR 29-SEP-1998; 98US-0102207P.
PR 29-SEP-1998; 98US-0102240P.
PR 29-SEP-1998; 98US-0102330P.
PR 29-SEP-1998; 98US-0102331P.
PR 30-SEP-1998; 98US-0102487P.
PR 30-SEP-1998; 98US-0102570P.
PR 30-SEP-1998; 98US-0102571P.
PR 01-OCT-1998; 98US-0102684P.
PR 01-OCT-1998; 98US-0102687P.
PR 02-OCT-1998; 98US-0102965P.
PR 06-OCT-1998; 98US-0103258P.
PR 06-OCT-1998; 98US-0103449P.

Query Match 98.7%; Score 1101; DB 6; Length 382;
Best Local Similarity 96.3%; Pred. No. 5.2e-103;
Matches 206; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

```
Qy          1 ATGRLLS--------GQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 52
              |||||||        ||||||||||||||||||||||||||||||||||||||||||||
Db         22 ATGRLLSASDLDLRGGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 81

Qy         53 ESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVD 112
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 ESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVD 141

Qy        113 EPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGE 172
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 EPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGE 201

Qy        173 ETELTTPVLPEETQEEDAKKTFKESREAALNLAY 206
              |||||||||||||||||||||||||||||||||
Db        202 ETELTTPVLPEETQEEDAKKTFKESREAALNLAY 235
```

RESULT 33
ABU91668
ID ABU91668 standard; protein; 382 AA.
XX
AC ABU91668;
XX
DT 11-AUG-2003 (first entry)
XX
DE Novel human secreted and transmembrane protein PRO234.
XX
KW Human; gene therapy; chromosome identification; tissue typing.
XX
OS Homo sapiens.
XX
PN US2003027277-A1.
XX
PD 06-FEB-2003.
XX
PF 21-JUN-2002; 2002US-00176985.
XX
PR 18-SEP-1997; 97US-0059263P.
PR 18-SEP-1997; 97US-0059266P.
PR 17-OCT-1997; 97US-0062250P.
PR 21-OCT-1997; 97US-0063486P.
PR 24-OCT-1997; 97US-0063120P.
PR 24-OCT-1997; 97US-0063121P.
PR 28-OCT-1997; 97US-0063540P.
PR 28-OCT-1997; 97US-0063541P.
PR 28-OCT-1997; 97US-0063544P.
PR 28-OCT-1997; 97US-0063564P.
PR 29-OCT-1997; 97US-0063734P.
PR 31-OCT-1997; 97US-0063870P.
PR 31-OCT-1997; 97US-0064103P.
PR 13-NOV-1997; 97US-0065311P.
PR 21-NOV-1997; 97US-0066120P.
PR 24-NOV-1997; 97US-0066466P.
PR 24-NOV-1997; 97US-0066772P.
PR 11-DEC-1997; 97US-0069335P.
PR 12-DEC-1997; 97US-0069425P.
PR 17-DEC-1997; 97US-0069870P.
PR 18-DEC-1997; 97US-0068017P.
PR 10-MAR-1998; 98US-0077450P.
PR 11-MAR-1998; 98US-0077632P.
PR 11-MAR-1998; 98US-0077649P.
PR 20-MAR-1998; 98US-0078886P.
PR 20-MAR-1998; 98US-0078939P.
PR 27-MAR-1998; 98US-0079664P.
PR 27-MAR-1998; 98US-0079786P.
PR 31-MAR-1998; 98US-0080107P.
PR 31-MAR-1998; 98US-0080194P.
PR 01-APR-1998; 98US-0080327P.
PR 01-APR-1998; 98US-0080333P.
PR 08-APR-1998; 98US-0081049P.
PR 08-APR-1998; 98US-0081070P.
PR 09-APR-1998; 98US-0081195P.
PR 15-APR-1998; 98US-0081838P.
PR 21-APR-1998; 98US-0082568P.
PR 21-APR-1998; 98US-0082569P.
PR 22-APR-1998; 98US-0082704P.
PR 22-APR-1998; 98US-0082797P.

PR 28-APR-1998; 98US-0083322P.
PR 29-APR-1998; 98US-0083495P.
PR 29-APR-1998; 98US-0083496P.
PR 29-APR-1998; 98US-0083499P.
PR 29-APR-1998; 98US-0083559P.
PR 05-MAY-1998; 98US-0084366P.
PR 06-MAY-1998; 98US-0084414P.
PR 07-MAY-1998; 98US-0084639P.
PR 07-MAY-1998; 98US-0084640P.
PR 07-MAY-1998; 98US-0084643P.
PR 15-MAY-1998; 98US-0085579P.
PR 15-MAY-1998; 98US-0085580P.
PR 15-MAY-1998; 98US-0085582P.
PR 15-MAY-1998; 98US-0085700P.
PR 18-MAY-1998; 98US-0086023P.
PR 22-MAY-1998; 98US-0086392P.
PR 22-MAY-1998; 98US-0086486P.
PR 28-MAY-1998; 98US-0087098P.
PR 28-MAY-1998; 98US-0087208P.
PR 02-JUN-1998; 98US-0087609P.
PR 02-JUN-1998; 98US-0087759P.
PR 03-JUN-1998; 98US-0087827P.
PR 04-JUN-1998; 98US-0088025P.
PR 04-JUN-1998; 98US-0088028P.
PR 04-JUN-1998; 98US-0088029P.
PR 04-JUN-1998; 98US-0088033P.
PR 04-JUN-1998; 98US-0088326P.
PR 05-JUN-1998; 98US-0088167P.
PR 05-JUN-1998; 98US-0088202P.
PR 05-JUN-1998; 98US-0088212P.
PR 05-JUN-1998; 98US-0088217P.
PR 09-JUN-1998; 98US-0088655P.
PR 10-JUN-1998; 98US-0088722P.
PR 10-JUN-1998; 98US-0088738P.
PR 10-JUN-1998; 98US-0088740P.
PR 10-JUN-1998; 98US-0088811P.
PR 10-JUN-1998; 98US-0088824P.
PR 10-JUN-1998; 98US-0088825P.
PR 10-JUN-1998; 98US-0088826P.
PR 11-JUN-1998; 98US-0088861P.
PR 11-JUN-1998; 98US-0088863P.
PR 11-JUN-1998; 98US-0088876P.
PR 12-JUN-1998; 98US-0089090P.
PR 12-JUN-1998; 98US-0089105P.
PR 16-JUN-1998; 98US-0089512P.
PR 16-JUN-1998; 98US-0089514P.
PR 17-JUN-1998; 98US-0089538P.
PR 17-JUN-1998; 98US-0089598P.
PR 17-JUN-1998; 98US-0089653P.
PR 18-JUN-1998; 98US-0089908P.
PR 19-JUN-1998; 98US-0089952P.
PR 22-JUN-1998; 98US-0090246P.
PR 22-JUN-1998; 98US-0090252P.
PR 22-JUN-1998; 98US-0090254P.
PR 24-JUN-1998; 98US-0090429P.
PR 24-JUN-1998; 98US-0090435P.
PR 24-JUN-1998; 98US-0090444P.
PR 24-JUN-1998; 98US-0090461P.
PR 24-JUN-1998; 98US-0090535P.
PR 24-JUN-1998; 98US-0090540P.
PR 25-JUN-1998; 98US-0090676P.
PR 25-JUN-1998; 98US-0090678P.
PR 25-JUN-1998; 98US-0090688P.
PR 25-JUN-1998; 98US-0090690P.
PR 25-JUN-1998; 98US-0090694P.
PR 25-JUN-1998; 98US-0090695P.
PR 25-JUN-1998; 98US-0090696P.
PR 26-JUN-1998; 98US-00105413.
PR 26-JUN-1998; 98US-0090862P.
PR 26-JUN-1998; 98US-0090863P.
PR 26-JUN-1998; 98US-0091010P.
PR 01-JUL-1998; 98US-0091359P.
PR 01-JUL-1998; 98US-0091544P.
PR 02-JUL-1998; 98US-0091478P.
PR 02-JUL-1998; 98US-0091486P.
PR 02-JUL-1998; 98US-0091626P.
PR 02-JUL-1998; 98US-0091628P.
PR 02-JUL-1998; 98US-0091632P.
PR 24-JUL-1998; 98US-0094006P.
PR 04-AUG-1998; 98US-0095282P.
PR 10-AUG-1998; 98US-0095998P.
PR 10-AUG-1998; 98US-0096012P.
PR 17-AUG-1998; 98US-0096757P.
PR 17-AUG-1998; 98US-0096766P.
PR 17-AUG-1998; 98US-0096867P.
PR 17-AUG-1998; 98US-0096891P.
PR 17-AUG-1998; 98US-0096897P.
PR 18-AUG-1998; 98US-0096949P.
PR 18-AUG-1998; 98US-0096959P.
PR 18-AUG-1998; 98US-0097022P.
PR 26-AUG-1998; 98US-0097952P.
PR 26-AUG-1998; 98US-0097954P.
PR 26-AUG-1998; 98US-0097955P.
PR 26-AUG-1998; 98US-0097971P.
PR 26-AUG-1998; 98US-0097974P.
PR 26-AUG-1998; 98US-0098014P.
PR 01-SEP-1998; 98US-0098716P.
PR 01-SEP-1998; 98US-0098723P.
PR 02-SEP-1998; 98US-0098803P.
PR 02-SEP-1998; 98US-0098821P.
PR 02-SEP-1998; 98US-0098843P.
PR 09-SEP-1998; 98US-0099602P.
PR 10-SEP-1998; 98US-0099741P.
PR 10-SEP-1998; 98US-0099754P.
PR 10-SEP-1998; 98US-0099763P.
PR 10-SEP-1998; 98US-0099812P.
PR 15-SEP-1998; 98US-0100388P.
PR 16-SEP-1998; 98US-0100662P.
PR 16-SEP-1998; 98US-0100664P.
PR 16-SEP-1998; 98US-0101751P.
PR 16-SEP-1998; 98WO-US019330.
PR 17-SEP-1998; 98US-0100683P.
PR 17-SEP-1998; 98US-0100684P.
PR 17-SEP-1998; 98US-0100919P.
PR 17-SEP-1998; 98US-0100930P.
PR 18-SEP-1998; 98US-0100849P.
PR 18-SEP-1998; 98US-0101014P.
PR 18-SEP-1998; 98US-0101068P.
PR 23-SEP-1998; 98US-0101471P.
PR 23-SEP-1998; 98US-0101472P.
PR 23-SEP-1998; 98US-0101475P.
PR 23-SEP-1998; 98US-0101477P.
PR 24-SEP-1998; 98US-0101738P.
PR 24-SEP-1998; 98US-0101739P.
PR 24-SEP-1998; 98US-0101743P.
PR 24-SEP-1998; 98US-0101922P.
PR 25-SEP-1998; 98US-0101786P.
PR 29-SEP-1998; 98US-0102207P.
PR 29-SEP-1998; 98US-0102240P.
PR 29-SEP-1998; 98US-0102330P.
PR 29-SEP-1998; 98US-0102331P.
PR 30-SEP-1998; 98US-0102487P.
PR 30-SEP-1998; 98US-0102570P.
PR 30-SEP-1998; 98US-0102571P.
PR 01-OCT-1998; 98US-0102684P.
PR 01-OCT-1998; 98US-0102687P.
PR 02-OCT-1998; 98US-0102965P.
PR 06-OCT-1998; 98US-0103258P.
PR 06-OCT-1998; 98US-0103449P.
PR 07-OCT-1998; 98US-00168978.
PR 07-OCT-1998; 98US-0103395P.
PR 07-OCT-1998; 98US-0103401P.

Query Match 98.7%; Score 1101; DB 6; Length 382;
Best Local Similarity 96.3%; Pred. No. 5.2e-103;
Matches 206; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

```
Qy          1 ATGRLLS--------GQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 52
              |||||||        ||||||||||||||||||||||||||||||||||||||||||||
Db         22 ATGRLLSASDLDLRGGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 81

Qy         53 ESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVD 112
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 ESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVD 141

Qy        113 EPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGE 172
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 EPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGE 201

Qy        173 ETELTTPVLPEETQEEDAKKTFKESREAALNLAY 206
              |||||||||||||||||||||||||||||||||
Db        202 ETELTTPVLPEETQEEDAKKTFKESREAALNLAY 235
```

RESULT 34
ABU71914
ID ABU71914 standard; protein; 382 AA.
XX
AC ABU71914;
XX
DT 12-JUN-2003 (first entry)
XX
DE Human secreted/transmembrane protein PRO234.
XX
KW Human; secreted protein; transmembrane protein; PRO; gene therapy;
KW chromosome identification; chromosome marker.
XX
OS Homo sapiens.
XX
PN US2003003530-A1.
XX
PD 02-JAN-2003.
XX
PF 11-JUL-2001; 2001US-00904011.
XX
PR 17-SEP-1997; 97US-0059113P.
PR 17-SEP-1997; 97US-0059115P.
PR 17-SEP-1997; 97US-0059117P.
PR 17-SEP-1997; 97US-0059119P.
PR 17-SEP-1997; 97US-0059121P.
PR 17-SEP-1997; 97US-0059122P.
PR 17-SEP-1997; 97US-0059184P.
PR 18-SEP-1997; 97US-0059263P.
PR 18-SEP-1997; 97US-0059266P.
PR 15-OCT-1997; 97US-0062125P.
PR 17-OCT-1997; 97US-0062285P.
PR 17-OCT-1997; 97US-0062287P.
PR 21-OCT-1997; 97US-0063486P.
PR 24-OCT-1997; 97US-0062814P.
PR 24-OCT-1997; 97US-0062816P.
PR 24-OCT-1997; 97US-0063045P.
PR 24-OCT-1997; 97US-0063120P.
PR 24-OCT-1997; 97US-0063121P.
PR 24-OCT-1997; 97US-0063127P.
PR 24-OCT-1997; 97US-0063128P.
PR 27-OCT-1997; 97US-0063327P.
PR 27-OCT-1997; 97US-0063329P.
PR 28-OCT-1997; 97US-0063541P.
PR 28-OCT-1997; 97US-0063542P.
PR 28-OCT-1997; 97US-0063544P.
PR 28-OCT-1997; 97US-0063549P.
PR 28-OCT-1997; 97US-0063550P.
PR 28-OCT-1997; 97US-0063564P.
PR 29-OCT-1997; 97US-0063435P.
PR 29-OCT-1997; 97US-0063704P.
PR 29-OCT-1997; 97US-0063732P.
PR 29-OCT-1997; 97US-0063734P.
PR 29-OCT-1997; 97US-0063735P.
PR 29-OCT-1997; 97US-0063738P.
PR 29-OCT-1997; 97US-0064215P.
PR 31-OCT-1997; 97US-0063870P.
PR 31-OCT-1997; 97US-0064103P.
PR 03-NOV-1997; 97US-0064248P.
PR 07-NOV-1997; 97US-0064809P.
PR 12-NOV-1997; 97US-0065186P.
PR 17-NOV-1997; 97US-0065846P.
PR 18-NOV-1997; 97US-0065693P.
PR 21-NOV-1997; 97US-0066120P.
PR 21-NOV-1997; 97US-0066364P.
PR 24-NOV-1997; 97US-0066453P.
PR 24-NOV-1997; 97US-0066466P.
PR 24-NOV-1997; 97US-0066511P.
PR 24-NOV-1997; 97US-0066770P.
PR 24-NOV-1997; 97US-0066772P.
PR 10-SEP-1998; 98WO-US018824.
PR 14-SEP-1998; 98WO-US019177.
PR 16-SEP-1998; 98WO-US019330.
PR 17-SEP-1998; 98WO-US019437.
PR 01-DEC-1998; 98WO-US025108.
PR 08-SEP-1999; 99WO-US020594.
PR 13-SEP-1999; 99WO-US020944.
PR 15-SEP-1999; 99WO-US021090.
PR 15-SEP-1999; 99WO-US021547.
PR 05-OCT-1999; 99WO-US023089.
PR 29-NOV-1999; 99WO-US028214.
PR 30-NOV-1999; 99WO-US028313.
PR 01-DEC-1999; 99WO-US028301.
PR 02-DEC-1999; 99WO-US028564.
PR 02-DEC-1999; 99WO-US028565.
PR 16-DEC-1999; 99WO-US030095.
PR 20-DEC-1999; 99WO-US030911.
PR 20-DEC-1999; 99WO-US030999.
PR 05-JAN-2000; 2000WO-US000219.
PR 11-FEB-2000; 2000WO-US003565.
PR 22-FEB-2000; 2000WO-US004414.
PR 24-FEB-2000; 2000WO-US005004.
PR 02-MAR-2000; 2000WO-US005841.
PR 20-MAR-2000; 2000WO-US007377.
PR 30-MAR-2000; 2000WO-US008439.
PR 22-MAY-2000; 2000WO-US014042.
PR 02-JUN-2000; 2000WO-US015264.
PR 28-JUL-2000; 2000WO-US020710.
PR 24-AUG-2000; 2000WO-US023328.
PR 18-SEP-2000; 2000US-00665350.
XX
PA (GETH) GENENTECH INC.
XX
PI Ashkenazi A, Botstein D, Desnoyers L, Eaton DL, Ferrara N;
PI Filvaroff E, Fong S, Gao W, Gerber H, Gerritsen ME, Goddard A;
PI Godowski PJ, Grimaldi JC, Gurney AL, Hillan KJ, Kljavin IJ;
PI Mather JP, Pan J, Paoni NF, Roy MA, Stewart TA, Tumas D;
PI Williams PM, Wood WI;
XX
DR WPI; 2003-329602/31.
DR N-PSDB; ACA60112.
XX
PT New transmembrane polypeptides and nucleic acids encoding the
PT polypeptides, useful in gene therapy, in chromosome identification, as
PT chromosome markers, in generating probes and in tissue typing.
XX
PS Claim 12; Fig 50; 484pp; English.
XX
CC The invention relates to an isolated nucleic acid with at least 80%
CC nucleic acid sequence identity to a nucleotide sequence encoding one of
CC 61 secreted/transmembrane polypeptides, or PRO polypeptides or encoding a
CC PRO protein extracellular domain. Also included are a vector comprising
CC the PRO nucleic acid, a host cell comprising the vector, producing a PRO
CC polypeptide (by culturing the host cell for the expression of the PRO
CC polypeptide, and recovering the PRO polypeptide from the cell culture),
CC an isolated PRO polypeptide (having at least 80% sequence identity to: (
CC a) an amino acid sequence selected from the 61 PRO proteins; (b) an amino
CC acid sequence encoded by a nucleic acid molecule deposited with an ATCC

```
CC   number (detailed in the specification); or (c) an extracellular domain of
CC   a PRO polypeptide or to a PRO polypeptide lacking its associated signal
CC   peptide), a chimaeric molecule comprising a PRO polypeptide of fused to a
CC   heterologous amino acid sequence, an anti-PRO antibody, detecting a
CC   PRO245 or PRO1868 in a sample suspected of containing the polypeptide,
CC   linking a bioactive molecule to a cell expressing a PRO245 or PRO1868 and
CC   modulating at least one biological activity of a cell expressing a PRO245
CC   or PRO1868. Nucleic acids which encode PRO can be used to generate either
CC   transgenic animals or knock-out animals which may be used in the
CC   development and screening of therapeutically useful reagents. The nucleic
CC   acids may also be used in gene therapy, in chromosome identification, as
CC   chromosome markers, or in generating probes. The PRO polypeptides are
CC   useful as molecular markers for protein electrophoresis, and the isolated
CC   nucleic acids may be used for recombinantly expressing those markers. The
CC   PRO polypeptides and nucleic acids may also be used in tissue typing.
CC   Anti-PRO antibodies are useful in diagnostic assays for PRO, and in
CC   affinity purification of PRO from recombinant cell culture or natural
CC   sources. The present sequence represents a PRO protein
XX
SQ   Sequence 382 AA;

  Query Match             98.7%;  Score 1101;  DB 6;  Length 382;
  Best Local Similarity   96.3%;  Pred. No. 5.2e-103;
  Matches  206;  Conservative    0;  Mismatches    0;  Indels    8;  Gaps    1;

Qy          1 ATGRLLS--------GQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 52
              |||||||        |||||||||||||||||||||||||||||||||||||||||||||
Db         22 ATGRLLSASDLDLRGGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 81

Qy         53 ESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVD 112
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 ESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVD 141

Qy        113 EPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGE 172
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 EPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGE 201

Qy        173 ETELTTPVLPEETQEEDAKKTFKESREAALNLAY 206
              ||||||||||||||||||||||||||||||||||
Db        202 ETELTTPVLPEETQEEDAKKTFKESREAALNLAY 235


RESULT 35
ABU89361
ID   ABU89361 standard; protein; 382 AA.
XX
AC   ABU89361;
XX
DT   09-JUL-2003  (first entry)
XX
DE   Human PRO polypeptide #10.
XX
KW   Human; PRO polypeptide; secreted protein; transmembrane protein;
KW   chromosome mapping; gene mapping; tumour; adrenal; lung; colon; breast;
KW   prostate; rectal; cervical; liver; cancer; TNF-alpha;
KW   tumour necrosis factor-alpha; proliferation; differentiation;
KW   chondrocyte cell; bone disorder; cartilage disorder; sports injury;
KW   arthritis; cytostatic; antiarthritic; osteopathic.
XX
OS   Homo sapiens.
XX
PN   US2003036141-A1.
XX
PD   20-FEB-2003.
XX
PF   01-JUL-2002; 2002US-00187597.
XX
PR   18-SEP-1997;   97US-0059263P.
PR   18-SEP-1997;   97US-0059266P.
PR   17-OCT-1997;   97US-0062250P.
PR   21-OCT-1997;   97US-0063486P.
PR   24-OCT-1997;   97US-0063120P.
PR   24-OCT-1997;   97US-0063121P.
PR   28-OCT-1997;   97US-0063540P.
PR   28-OCT-1997;   97US-0063541P.
PR   28-OCT-1997;   97US-0063544P.
PR   28-OCT-1997;   97US-0063564P.
PR   29-OCT-1997;   97US-0063734P.
PR   31-OCT-1997;   97US-0063870P.
PR   31-OCT-1997;   97US-0064103P.
PR   13-NOV-1997;   97US-0065311P.
PR   21-NOV-1997;   97US-0066120P.
PR   24-NOV-1997;   97US-0066466P.
PR   24-NOV-1997;   97US-0066772P.
PR   11-DEC-1997;   97US-0069335P.
PR   12-DEC-1997;   97US-0069425P.
PR   17-DEC-1997;   97US-0069870P.
PR   18-DEC-1997;   97US-0068017P.
PR   10-MAR-1998;   98US-0077450P.
PR   11-MAR-1998;   98US-0077632P.
PR   11-MAR-1998;   98US-0077649P.
PR   20-MAR-1998;   98US-0078886P.
PR   20-MAR-1998;   98US-0078939P.
PR   27-MAR-1998;   98US-0079664P.
PR   27-MAR-1998;   98US-0079786P.
PR   31-MAR-1998;   98US-0080107P.
PR   31-MAR-1998;   98US-0080194P.
PR   01-APR-1998;   98US-0080327P.
PR   01-APR-1998;   98US-0080333P.
PR   08-APR-1998;   98US-0081049P.
PR   08-APR-1998;   98US-0081070P.
PR   09-APR-1998;   98US-0081195P.
PR   15-APR-1998;   98US-0081838P.
PR   21-APR-1998;   98US-0082568P.
PR   21-APR-1998;   98US-0082569P.
PR   22-APR-1998;   98US-0082704P.
PR   22-APR-1998;   98US-0082797P.
PR   28-APR-1998;   98US-0083322P.
PR   29-APR-1998;   98US-0083495P.
PR   29-APR-1998;   98US-0083496P.
PR   29-APR-1998;   98US-0083499P.
PR   29-APR-1998;   98US-0083559P.
PR   05-MAY-1998;   98US-0084366P.
PR   06-MAY-1998;   98US-0084414P.
PR   07-MAY-1998;   98US-0084639P.
PR   07-MAY-1998;   98US-0084640P.
PR   07-MAY-1998;   98US-0084643P.
PR   15-MAY-1998;   98US-0085579P.
PR   15-MAY-1998;   98US-0085580P.
PR   15-MAY-1998;   98US-0085582P.
PR   15-MAY-1998;   98US-0085700P.
PR   18-MAY-1998;   98US-0086023P.
PR   22-MAY-1998;   98US-0086392P.
PR   22-MAY-1998;   98US-0086486P.
PR   28-MAY-1998;   98US-0087098P.
PR   28-MAY-1998;   98US-0087208P.
PR   02-JUN-1998;   98US-0087609P.
PR   02-JUN-1998;   98US-0087759P.
PR   03-JUN-1998;   98US-0087827P.
PR   04-JUN-1998;   98US-0088025P.
PR   04-JUN-1998;   98US-0088028P.
PR   04-JUN-1998;   98US-0088029P.
PR   04-JUN-1998;   98US-0088033P.
PR   04-JUN-1998;   98US-0088326P.
PR   05-JUN-1998;   98US-0088167P.
PR   05-JUN-1998;   98US-0088202P.
PR   05-JUN-1998;   98US-0088212P.
PR   05-JUN-1998;   98US-0088217P.
PR   09-JUN-1998;   98US-0088655P.
PR   10-JUN-1998;   98US-0088722P.
PR   10-JUN-1998;   98US-0088738P.
PR   10-JUN-1998;   98US-0088740P.
PR   10-JUN-1998;   98US-0088811P.
PR   10-JUN-1998;   98US-0088824P.
PR   10-JUN-1998;   98US-0088825P.
PR   10-JUN-1998;   98US-0088826P.
```

PR 11-JUN-1998; 98US-0088861P.
PR 11-JUN-1998; 98US-0088863P.
PR 11-JUN-1998; 98US-0088876P.
PR 12-JUN-1998; 98US-0089090P.
PR 12-JUN-1998; 98US-0089105P.
PR 16-JUN-1998; 98US-0089512P.
PR 16-JUN-1998; 98US-0089514P.
PR 17-JUN-1998; 98US-0089538P.
PR 17-JUN-1998; 98US-0089598P.
PR 17-JUN-1998; 98US-0089653P.
PR 18-JUN-1998; 98US-0089908P.
PR 19-JUN-1998; 98US-0089952P.
PR 22-JUN-1998; 98US-0090246P.
PR 22-JUN-1998; 98US-0090252P.
PR 22-JUN-1998; 98US-0090254P.
PR 24-JUN-1998; 98US-0090429P.
PR 24-JUN-1998; 98US-0090435P.
PR 24-JUN-1998; 98US-0090444P.
PR 24-JUN-1998; 98US-0090461P.
PR 24-JUN-1998; 98US-0090535P.
PR 24-JUN-1998; 98US-0090540P.
PR 25-JUN-1998; 98US-0090676P.
PR 25-JUN-1998; 98US-0090678P.
PR 25-JUN-1998; 98US-0090688P.
PR 25-JUN-1998; 98US-0090690P.
PR 25-JUN-1998; 98US-0090694P.
PR 25-JUN-1998; 98US-0090695P.
PR 25-JUN-1998; 98US-0090696P.
PR 26-JUN-1998; 98US-00105413.
PR 26-JUN-1998; 98US-0090862P.
PR 26-JUN-1998; 98US-0090863P.
PR 26-JUN-1998; 98US-0091010P.
PR 01-JUL-1998; 98US-0091359P.
PR 01-JUL-1998; 98US-0091544P.
PR 02-JUL-1998; 98US-0091478P.
PR 02-JUL-1998; 98US-0091486P.
PR 02-JUL-1998; 98US-0091626P.
PR 02-JUL-1998; 98US-0091628P.
PR 02-JUL-1998; 98US-0091632P.
PR 24-JUL-1998; 98US-0094006P.
PR 04-AUG-1998; 98US-0095282P.
PR 10-AUG-1998; 98US-0095998P.
PR 10-AUG-1998; 98US-0096012P.
PR 17-AUG-1998; 98US-0096757P.
PR 17-AUG-1998; 98US-0096766P.
PR 17-AUG-1998; 98US-0096867P.
PR 17-AUG-1998; 98US-0096891P.
PR 17-AUG-1998; 98US-0096897P.
PR 18-AUG-1998; 98US-0096949P.
PR 18-AUG-1998; 98US-0096959P.
PR 18-AUG-1998; 98US-0097022P.
PR 26-AUG-1998; 98US-0097952P.
PR 26-AUG-1998; 98US-0097954P.
PR 26-AUG-1998; 98US-0097955P.
PR 26-AUG-1998; 98US-0097971P.
PR 26-AUG-1998; 98US-0097974P.
PR 26-AUG-1998; 98US-0098014P.
PR 01-SEP-1998; 98US-0098716P.
PR 01-SEP-1998; 98US-0098723P.
PR 02-SEP-1998; 98US-0098803P.
PR 02-SEP-1998; 98US-0098821P.
PR 02-SEP-1998; 98US-0098843P.
PR 09-SEP-1998; 98US-0099602P.
PR 10-SEP-1998; 98US-0099741P.
PR 10-SEP-1998; 98US-0099754P.
PR 10-SEP-1998; 98US-0099763P.
PR 10-SEP-1998; 98US-0099812P.
PR 15-SEP-1998; 98US-0100388P.
PR 16-SEP-1998; 98US-0100662P.
PR 16-SEP-1998; 98US-0100664P.
PR 16-SEP-1998; 98US-0101751P.
PR 16-SEP-1998; 98WO-US019330.
PR 17-SEP-1998; 98US-0100683P.
PR 17-SEP-1998; 98US-0100684P.
PR 17-SEP-1998; 98US-0100919P.
PR 17-SEP-1998; 98US-0100930P.
PR 18-SEP-1998; 98US-0100849P.
PR 18-SEP-1998; 98US-0101014P.
PR 18-SEP-1998; 98US-0101068P.
PR 23-SEP-1998; 98US-0101471P.
PR 23-SEP-1998; 98US-0101472P.
PR 23-SEP-1998; 98US-0101475P.
PR 23-SEP-1998; 98US-0101477P.
PR 24-SEP-1998; 98US-0101738P.
PR 24-SEP-1998; 98US-0101739P.
PR 24-SEP-1998; 98US-0101743P.
PR 24-SEP-1998; 98US-0101922P.
PR 25-SEP-1998; 98US-0101786P.
PR 29-SEP-1998; 98US-0102207P.
PR 29-SEP-1998; 98US-0102240P.
PR 29-SEP-1998; 98US-0102330P.
PR 29-SEP-1998; 98US-0102331P.
PR 30-SEP-1998; 98US-0102487P.
PR 30-SEP-1998; 98US-0102570P.
PR 30-SEP-1998; 98US-0102571P.
PR 01-OCT-1998; 98US-0102684P.
PR 01-OCT-1998; 98US-0102687P.
PR 02-OCT-1998; 98US-0102965P.

Query Match 98.7%; Score 1101; DB 6; Length 382;
Best Local Similarity 96.3%; Pred. No. 5.2e-103;
Matches 206; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

```
Qy          1 ATGRLLS--------GQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 52
              |||||||        ||||||||||||||||||||||||||||||||||||||||||||
Db         22 ATGRLLSASDLDLRGGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 81

Qy         53 ESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVD 112
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 ESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVD 141

Qy        113 EPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGE 172
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 EPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGE 201

Qy        173 ETELTTPVLPEETQEEDAKKTFKESREAALNLAY 206
              |||||||||||||||||||||||||||||||||
Db        202 ETELTTPVLPEETQEEDAKKTFKESREAALNLAY 235
```

RESULT 36
ABU86202
ID ABU86202 standard; protein; 382 AA.
XX
AC ABU86202;
XX
DT 01-JUL-2003 (first entry)
XX
DE Human secreted/transmembrane protein (PRO) #10.
XX
KW Human; immunogen; secreted protein; transmembrane protein; PRO; tumour;
KW proliferation; differentiation; chondrocyte cells;
KW tumour necrosis factor-alpha; TNF-alpha; blood; gene therapy.
XX
OS Homo sapiens.
XX
PN US2003036146-A1.
XX
PD 20-FEB-2003.
XX
PF 02-JUL-2002; 2002US-00187603.
XX
PR 26-JUN-1998; 98US-00105413.
PR 16-SEP-1998; 98WO-US019330.
PR 07-OCT-1998; 98US-00168978.
PR 07-OCT-1998; 98WO-US021141.

PR 06-NOV-1998; 98US-00187368.
PR 01-DEC-1998; 98WO-US025108.
PR 07-DEC-1998; 98US-00202054.
PR 03-MAR-1999; 99US-00254311.
PR 08-MAR-1999; 99WO-US005028.
PR 14-MAY-1999; 99US-00311832.
PR 14-MAY-1999; 99WO-US010733.
PR 02-JUN-1999; 99WO-US012252.
PR 25-AUG-1999; 99US-00380137.
PR 25-AUG-1999; 99US-00380138.
PR 25-AUG-1999; 99US-00380139.
PR 25-AUG-1999; 99US-00380142.
PR 01-SEP-1999; 99WO-US020111.
PR 15-SEP-1999; 99WO-US021090.
PR 18-OCT-1999; 99US-00403297.
PR 12-NOV-1999; 99US-00423844.
PR 01-DEC-1999; 99WO-US028301.
PR 02-DEC-1999; 99WO-US028551.
PR 30-DEC-1999; 99WO-US031274.
PR 05-JAN-2000; 2000WO-US000219.
PR 18-FEB-2000; 2000WO-US004341.
PR 18-FEB-2000; 2000WO-US004342.
PR 22-FEB-2000; 2000WO-US004414.
PR 24-FEB-2000; 2000WO-US005004.
PR 01-MAR-2000; 2000WO-US005601.
PR 02-MAR-2000; 2000WO-US005841.
PR 15-MAR-2000; 2000WO-US006884.
PR 30-MAR-2000; 2000WO-US008439.
PR 17-MAY-2000; 2000WO-US013705.
PR 22-MAY-2000; 2000WO-US014042.
PR 30-MAY-2000; 2000WO-US014941.
PR 02-JUN-2000; 2000WO-US015264.
PR 28-JUL-2000; 2000WO-US020710.
PR 22-AUG-2000; 2000US-00644848.
PR 24-AUG-2000; 2000WO-US023328.
PR 18-SEP-2000; 2000US-00664610.
PR 18-SEP-2000; 2000US-00665350.
PR 08-NOV-2000; 2000US-00709238.
PR 08-NOV-2000; 2000WO-US030952.
PR 01-DEC-2000; 2000WO-US032678.
PR 20-DEC-2000; 2000US-00747259.
PR 20-DEC-2000; 2000WO-US034956.
PR 28-FEB-2001; 2001WO-US006520.
PR 22-MAR-2001; 2001US-00816744.
PR 10-MAY-2001; 2001US-00854208.
PR 10-MAY-2001; 2001US-00854280.
PR 25-MAY-2001; 2001US-00866028.
PR 01-JUN-2001; 2001WO-US017800.
PR 05-JUN-2001; 2001US-00874503.
PR 20-JUN-2001; 2001WO-US019692.
PR 29-JUN-2001; 2001WO-US021066.
PR 09-JUL-2001; 2001WO-US021735.
PR 18-JUL-2001; 2001US-00908827.
PR 30-JUL-2001; 2001US-00918585.
PR 06-AUG-2001; 2001US-00924419.
PR 13-AUG-2001; 2001US-00929404.
PR 16-AUG-2001; 2001US-00931836.
PR 28-AUG-2001; 2001US-00941992.
PR 29-AUG-2001; 2001WO-US027099.
PR 04-SEP-2001; 2001US-00946374.
PR 15-JAN-2002; 2002US-00052586.
XX
PA (GETH) GENENTECH INC.
XX
PI Baker KP, Chen J, Desnoyers L, Goddard A, Godowski PJ, Gurney AL;
PI Pan J, Smith V, Watanabe CK, Wood WI, Zhang Z;
XX
DR WPI; 2003-332034/31.
DR N-PSDB; ACA73394.
XX
PT Three hundred and five nucleic acids encoding PRO polypeptides, useful in
PT gene therapy, chromosome identification, tissue typing, and for detecting
PT the presence of tumor in a mammal.
XX
PS Claim 11; Fig 20; 707pp; English.
XX
CC The invention relates to three hundred and five nucleic acids encoding
CC PRO polypeptides (secreted and transmembrane), sequences 80% identical to
CC them, or encoding a PRO polypeptide lacking its associated signal peptide
CC or an extracellular domain of the PRO polypeptide, with or lacking its
CC associated signal peptide. Also included are the encoded PRO proteins,
CC PRO expression vectors, host cells transformed with the vector (used to
CC produce PRO proteins), a chimaeric molecule comprising the PRO
CC polypeptide fused to a heterologous amino acid sequence, an anti-PRO
CC antibody, a method for stimulating the release of tumor necrosis factor
CC alpha (TNF-alpha) from human blood (by contacting the blood with PRO1079,
CC PRO827, PRO791, PRO1131, PRO1316, PRO1183, PRO1343, PRO1760, PRO1567 or
CC PRO4333), a method for stimulating the proliferation or differentiation
CC of chondrocyte cells by contacting the cells with a PRO6029 polypeptide,
CC a method for detecting the presence of tumour in a mammal and an
CC oligonucleotide probe derived from any of the nucleotide sequences cited
CC above. The PRO polypeptide or anti-PRO antibody is useful for preparing a
CC medicament for treating a condition that is responsive to the PRO
CC polypeptide or anti-PRO antibody. The PRO nucleotide sequences are useful
CC as hybridisation probes in chromosome and gene mapping, or in generating
CC antisense RNA and DNA. PRO nucleic acids are also useful in preparing PRO
CC polypeptides, in assays to identify other proteins or molecules involved
CC in a binding reaction, to generate transgenic animals or knockout
CC animals, which in turn are useful in the development and screening of
CC therapeutically useful reagents, for chromosome identification, and
CC tissue typing. The PRO polypeptides and nucleic acid molecules are also
CC useful for detecting the presence of a tumour in a mammal, stimulating
CC proliferation or differentiation of chondrocyte cells, stimulating the
CC release of tumour necrosis factor-alpha from human blood, in gene
CC therapy, or as molecular weight markers for protein electrophoresis
CC purposes. The anti-PRO antibodies may be used in diagnostic assays for
CC PRO, or for the affinity purification of PRO from recombinant cell
CC culture or natural sources. The present sequence represents a PRO protein
XX
SQ Sequence 382 AA;

```
  Query Match             98.7%;  Score 1101;  DB 6;  Length 382;
  Best Local Similarity   96.3%;  Pred. No. 5.2e-103;
  Matches  206;  Conservative    0;  Mismatches    0;  Indels    8;  Gaps    1;

Qy          1 ATGRLLS--------GQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 52
              |||||||        ||||||||||||||||||||||||||||||||||||||||||||
Db         22 ATGRLLSASDLDLRGGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 81

Qy         53 ESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVD 112
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 ESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVD 141

Qy        113 EPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGE 172
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 EPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGE 201

Qy        173 ETELTTPVLPEETQEEDAKKTFKESREAALNLAY 206
              ||||||||||||||||||||||||||||||||||
Db        202 ETELTTPVLPEETQEEDAKKTFKESREAALNLAY 235
```

RESULT 37
ABU67415
ID ABU67415 standard; protein; 382 AA.
XX
AC ABU67415;
XX
DT 29-MAY-2003 (first entry)
XX
DE Human secreted/transmembrane protein (PRO) #10.
XX
KW Human; secreted and transmembrane protein; PRO; TNF-alpha;
KW tumour necrosis factor alpha; chondrocyte cell; tumour; gene therapy;
KW tissue typing.
XX

OS Homo sapiens.
XX
PN US2003036162-A1.
XX
PD 20-FEB-2003.
XX
PF 12-JUL-2002; 2002US-00194423.
XX
PR 26-JUN-1998; 98US-00105413.
PR 16-SEP-1998; 98WO-US019330.
PR 07-OCT-1998; 98US-00168978.
PR 07-OCT-1998; 98WO-US021141.
PR 06-NOV-1998; 98US-00187368.
PR 01-DEC-1998; 98WO-US025108.
PR 07-DEC-1998; 98US-00202054.
PR 03-MAR-1999; 99US-00254311.
PR 08-MAR-1999; 99WO-US005028.
PR 14-MAY-1999; 99US-00311832.
PR 14-MAY-1999; 99WO-US010733.
PR 02-JUN-1999; 99WO-US012252.
PR 25-AUG-1999; 99US-00380137.
PR 25-AUG-1999; 99US-00380138.
PR 25-AUG-1999; 99US-00380139.
PR 25-AUG-1999; 99US-00380142.
PR 01-SEP-1999; 99WO-US020111.
PR 15-SEP-1999; 99WO-US021090.
PR 18-OCT-1999; 99US-00403297.
PR 12-NOV-1999; 99US-00423844.
PR 01-DEC-1999; 99WO-US028301.
PR 02-DEC-1999; 99WO-US028551.
PR 30-DEC-1999; 99WO-US031274.
PR 05-JAN-2000; 2000WO-US000219.
PR 18-FEB-2000; 2000WO-US004341.
PR 18-FEB-2000; 2000WO-US004342.
PR 22-FEB-2000; 2000WO-US004414.
PR 24-FEB-2000; 2000WO-US005004.
PR 01-MAR-2000; 2000WO-US005601.
PR 02-MAR-2000; 2000WO-US005841.
PR 15-MAR-2000; 2000WO-US006884.
PR 30-MAR-2000; 2000WO-US008439.
PR 17-MAY-2000; 2000WO-US013705.
PR 22-MAY-2000; 2000WO-US014042.
PR 30-MAY-2000; 2000WO-US014941.
PR 02-JUN-2000; 2000WO-US015264.
PR 28-JUL-2000; 2000WO-US020710.
PR 22-AUG-2000; 2000US-00644848.
PR 24-AUG-2000; 2000WO-US023328.
PR 18-SEP-2000; 2000US-00664610.
PR 18-SEP-2000; 2000US-00665350.
PR 08-NOV-2000; 2000US-00709238.
PR 08-NOV-2000; 2000WO-US030952.
PR 01-DEC-2000; 2000WO-US032678.
PR 20-DEC-2000; 2000US-00747259.
PR 20-DEC-2000; 2000WO-US034956.
PR 28-FEB-2001; 2001WO-US006520.
PR 22-MAR-2001; 2001US-00816744.
PR 10-MAY-2001; 2001US-00854208.
PR 10-MAY-2001; 2001US-00854280.
PR 25-MAY-2001; 2001US-00866028.
PR 01-JUN-2001; 2001WO-US017800.
PR 05-JUN-2001; 2001US-00874503.
PR 20-JUN-2001; 2001WO-US019692.
PR 29-JUN-2001; 2001WO-US021066.
PR 09-JUL-2001; 2001WO-US021735.
PR 18-JUL-2001; 2001US-00908827.
PR 30-JUL-2001; 2001US-00918585.
PR 06-AUG-2001; 2001US-00924419.
PR 13-AUG-2001; 2001US-00929404.
PR 16-AUG-2001; 2001US-00931836.
PR 28-AUG-2001; 2001US-00941992.
PR 29-AUG-2001; 2001WO-US027099.
PR 04-SEP-2001; 2001US-00946374.
PR 15-JAN-2002; 2002US-00052586.
XX
PA (GETH) GENENTECH INC.
XX
PI Baker KP, Chen J, Desnoyers L, Goddard A, Godowski PJ, Gurney AL;
PI Pan J, Smith V, Watanabe CK, Wood WI, Zhang Z;
XX
DR WPI; 2003-332039/31.
DR N-PSDB; ACA05709.
XX
PT New secreted and transmembrane PRO polypeptides and nucleic acids, useful
PT in gene therapy, in chromosome and gene mapping, as chromosome markers,
PT in tissue typing, and in chromosome identification.
XX
PS Claim 11; Fig 20; 706pp; English.
XX
CC The invention discloses human nucleic acids encoding secreted and
CC transmembrane (PRO) polypeptides. Also disclosed is an antibody that
CC specifically binds to the PRO polypeptide, a method for stimulating the
CC release of tumour necrosis factor alpha (TNF-alpha) from human blood by
CC contacting the blood a PRO polypeptide, a method for stimulating the
CC proliferation or differentiation of chondrocyte cells by contacting the
CC cells with a PRO polypeptide, a method for detecting the presence of a
CC tumour in a mammal and an oligonucleotide probe derived from any of the
CC PRO nucleotide sequences. The nucleotide sequences are useful as probes,
CC in chromosome and gene mapping, in generating antisense RNA and DNA, in
CC preparing PRO polypeptides by recombinant techniques and in gene therapy
CC (e.g. for replacement of defective gene). The PRO polypeptides are useful
CC as molecular weight markers for protein electrophoresis purposes, for
CC chromosome identification, as chromosome markers, as therapeutic agents,
CC for stimulating the release of TNF-alpha from human blood, for
CC stimulating the proliferation or differentiation of chondrocytes and
CC detecting the presence of a tumour. The PRO polypeptides and nucleic
CC acids may also be used diagnostically for tissue typing. The sequences
CC presented in ABU67406-ABU67710 are the PRO polypeptides of the invention
XX
SQ Sequence 382 AA;

Query Match 98.7%; Score 1101; DB 6; Length 382;
Best Local Similarity 96.3%; Pred. No. 5.2e-103;
Matches 206; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

```
Qy          1 ATGRLLS--------GQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 52
              |||||||        ||||||||||||||||||||||||||||||||||||||||||||
Db         22 ATGRLLSASDLDLRGGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 81

Qy         53 ESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVD 112
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 ESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVD 141

Qy        113 EPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGE 172
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 EPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGE 201

Qy        173 ETELTTPVLPEETQEEDAKKTFKESREAALNLAY 206
              ||||||||||||||||||||||||||||||||||
Db        202 ETELTTPVLPEETQEEDAKKTFKESREAALNLAY 235
```

RESULT 38
ABU80443
ID ABU80443 standard; protein; 382 AA.
XX
AC ABU80443;
XX
DT 23-JUN-2003 (first entry)
XX
DE Human PRO protein #10.
XX
KW Human; tumour; adrenal; lung; colon; breast; prostate; rectal; cervical;
KW liver; PRO; gene therapy.
XX
OS Homo sapiens.
XX

PN US2003036137-A1.
XX
PD 20-FEB-2003.
XX
PF 27-JUN-2002; 2002US-00184640.
XX
PR 26-JUN-1998; 98US-00105413.
PR 16-SEP-1998; 98WO-US019330.
PR 07-OCT-1998; 98US-00168978.
PR 07-OCT-1998; 98WO-US021141.
PR 06-NOV-1998; 98US-00187368.
PR 01-DEC-1998; 98WO-US025108.
PR 07-DEC-1998; 98US-00202054.
PR 03-MAR-1999; 99US-00254311.
PR 08-MAR-1999; 99WO-US005028.
PR 14-MAY-1999; 99US-00311832.
PR 14-MAY-1999; 99WO-US010733.
PR 02-JUN-1999; 99WO-US012252.
PR 25-AUG-1999; 99US-00380137.
PR 25-AUG-1999; 99US-00380138.
PR 25-AUG-1999; 99US-00380139.
PR 25-AUG-1999; 99US-00380142.
PR 01-SEP-1999; 99WO-US020111.
PR 15-SEP-1999; 99WO-US021090.
PR 18-OCT-1999; 99US-00403297.
PR 12-NOV-1999; 99US-00423844.
PR 01-DEC-1999; 99WO-US028301.
PR 02-DEC-1999; 99WO-US028551.
PR 30-DEC-1999; 99WO-US031274.
PR 05-JAN-2000; 2000WO-US000219.
PR 18-FEB-2000; 2000WO-US004341.
PR 18-FEB-2000; 2000WO-US004342.
PR 22-FEB-2000; 2000WO-US004414.
PR 24-FEB-2000; 2000WO-US005004.
PR 01-MAR-2000; 2000WO-US005601.
PR 02-MAR-2000; 2000WO-US005841.
PR 15-MAR-2000; 2000WO-US006884.
PR 30-MAR-2000; 2000WO-US008439.
PR 17-MAY-2000; 2000WO-US013705.
PR 22-MAY-2000; 2000WO-US014042.
PR 30-MAY-2000; 2000WO-US014941.
PR 02-JUN-2000; 2000WO-US015264.
PR 28-JUL-2000; 2000WO-US020710.
PR 22-AUG-2000; 2000US-00644848.
PR 24-AUG-2000; 2000WO-US023328.
PR 18-SEP-2000; 2000US-00664610.
PR 18-SEP-2000; 2000US-00665350.
PR 08-NOV-2000; 2000US-00709238.
PR 08-NOV-2000; 2000WO-US030952.
PR 01-DEC-2000; 2000WO-US032678.
PR 20-DEC-2000; 2000US-00747259.
PR 20-DEC-2000; 2000WO-US034956.
PR 28-FEB-2001; 2001WO-US006520.
PR 22-MAR-2001; 2001US-00816744.
PR 10-MAY-2001; 2001US-00854208.
PR 10-MAY-2001; 2001US-00854280.
PR 25-MAY-2001; 2001US-00866028.
PR 01-JUN-2001; 2001WO-US017800.
PR 05-JUN-2001; 2001US-00874503.
PR 20-JUN-2001; 2001WO-US019692.
PR 29-JUN-2001; 2001WO-US021066.
PR 09-JUL-2001; 2001WO-US021735.
PR 18-JUL-2001; 2001US-00908827.
PR 30-JUL-2001; 2001US-00918585.
PR 06-AUG-2001; 2001US-00924419.
PR 13-AUG-2001; 2001US-00929404.
PR 16-AUG-2001; 2001US-00931836.
PR 28-AUG-2001; 2001US-00941992.
PR 29-AUG-2001; 2001WO-US027099.
PR 04-SEP-2001; 2001US-00946374.
PR 15-JAN-2002; 2002US-00052586.
XX
PA (GETH) GENENTECH INC.
XX
PI Baker KP, Chen J, Desnoyers L, Goddard A, Godowski PJ, Gurney AL;
PI Pan J, Smith V, Watanabe CK, Wood WI, Zhang Z;
XX
DR WPI; 2003-342038/32.
DR N-PSDB; ACA66543.
XX
PT Three hundred and five nucleic acids encoding secreted and transmembrane
PT PRO polypeptides, useful for the diagnosis, prevention and/or treatment
PT of tumors, such as adrenal, lung, colon, breast, prostate, rectal,
PT cervical or liver tumors.
XX
PS Claim 11; Fig 20; 708pp; English.
XX
CC The invention relates to three hundred and five nucleic acids encoding
CC PRO polypeptides (secreted and transmembrane). Methods and compositions
CC of the present invention are useful for the diagnosis, prevention and/or
CC treatment of tumours, such as adrenal, lung, colon, breast, prostate,
CC rectal, cervical or liver tumours. The PRO polypeptides are also useful
CC as molecular weight markers, or for chromosome identification. The PRO
CC genes are useful as hybridisation probes, or for screening libraries of
CC human cDNA, genomic DNA or mRNA. The PRO genes may also be used in gene
CC therapy, particularly for replacing a defective gene. The present
CC sequence represents a human PRO polypeptide of the invention
XX
SQ Sequence 382 AA;

Query Match 98.7%; Score 1101; DB 6; Length 382;
Best Local Similarity 96.3%; Pred. No. 5.2e-103;
Matches 206; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

```
Qy          1 ATGRLLS--------GQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 52
              |||||||        ||||||||||||||||||||||||||||||||||||||||||||
Db         22 ATGRLLSASDLDLRGGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 81

Qy         53 ESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVD 112
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 ESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVD 141

Qy        113 EPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGE 172
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 EPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGE 201

Qy        173 ETELTTPVLPEETQEEDAKKTFKESREAALNLAY 206
              |||||||||||||||||||||||||||||||||
Db        202 ETELTTPVLPEETQEEDAKKTFKESREAALNLAY 235
```

RESULT 39
ABO01797
ID ABO01797 standard; protein; 382 AA.
XX
AC ABO01797;
XX
DT 07-AUG-2003 (first entry)
XX
DE Novel human secreted and transmembrane protein PRO234.
XX
KW Human; secreted and transmembrane protein; PRO; pharmaceutical;
KW diagnostic; biosensor; bioreactor; Parkinson's disease;
KW Alzheimer's disease; inflammation; nephritis; wound healing;
KW nerve repair; collateral blood vessel formation; cancer;
KW colorectal cancer; haemorrhage; rheumatoid arthritis; diabetes;
KW cirrhosis; fibrosis; restenosis; dermal fibrotic condition; keloid;
KW scarring; ischaemia; stroke; hypertension; heart attack; atherosclerosis;
KW infertility; gene therapy.
XX
OS Homo sapiens.
XX
PN US2002197671-A1.
XX
PD 26-DEC-2002.
XX

PF 17-JUL-2001; 2001US-00907824.
XX
PR 17-SEP-1997; 97US-0059113P.
PR 17-SEP-1997; 97US-0059115P.
PR 17-SEP-1997; 97US-0059117P.
PR 17-SEP-1997; 97US-0059119P.
PR 17-SEP-1997; 97US-0059121P.
PR 17-SEP-1997; 97US-0059122P.
PR 17-SEP-1997; 97US-0059184P.
PR 18-SEP-1997; 97US-0059263P.
PR 18-SEP-1997; 97US-0059266P.
PR 15-OCT-1997; 97US-0062125P.
PR 17-OCT-1997; 97US-0062285P.
PR 17-OCT-1997; 97US-0062287P.
PR 21-OCT-1997; 97US-0063486P.
PR 24-OCT-1997; 97US-0062814P.
PR 24-OCT-1997; 97US-0062816P.
PR 24-OCT-1997; 97US-0063045P.
PR 24-OCT-1997; 97US-0063120P.
PR 24-OCT-1997; 97US-0063121P.
PR 24-OCT-1997; 97US-0063127P.
PR 24-OCT-1997; 97US-0063128P.
PR 27-OCT-1997; 97US-0063327P.
PR 27-OCT-1997; 97US-0063329P.
PR 28-OCT-1997; 97US-0063541P.
PR 28-OCT-1997; 97US-0063542P.
PR 28-OCT-1997; 97US-0063544P.
PR 28-OCT-1997; 97US-0063549P.
PR 28-OCT-1997; 97US-0063550P.
PR 28-OCT-1997; 97US-0063564P.
PR 29-OCT-1997; 97US-0063435P.
PR 29-OCT-1997; 97US-0063704P.
PR 29-OCT-1997; 97US-0063732P.
PR 29-OCT-1997; 97US-0063734P.
PR 29-OCT-1997; 97US-0063735P.
PR 29-OCT-1997; 97US-0063738P.
PR 29-OCT-1997; 97US-0064215P.
PR 31-OCT-1997; 97US-0063870P.
PR 31-OCT-1997; 97US-0064103P.
PR 03-NOV-1997; 97US-0064248P.
PR 07-NOV-1997; 97US-0064809P.
PR 12-NOV-1997; 97US-0065186P.
PR 17-NOV-1997; 97US-0065846P.
PR 18-NOV-1997; 97US-0065693P.
PR 21-NOV-1997; 97US-0066120P.
PR 21-NOV-1997; 97US-0066364P.
PR 24-NOV-1997; 97US-0066453P.
PR 24-NOV-1997; 97US-0066466P.
PR 24-NOV-1997; 97US-0066511P.
PR 24-NOV-1997; 97US-0066770P.
PR 24-NOV-1997; 97US-0066772P.
PR 10-SEP-1998; 98WO-US018824.
PR 14-SEP-1998; 98WO-US019177.
PR 16-SEP-1998; 98WO-US019330.
PR 17-SEP-1998; 98WO-US019437.
PR 01-DEC-1998; 98WO-US025108.
PR 08-SEP-1999; 99WO-US020594.
PR 13-SEP-1999; 99WO-US020944.
PR 15-SEP-1999; 99WO-US021090.
PR 15-SEP-1999; 99WO-US021547.
PR 05-OCT-1999; 99WO-US023089.
PR 29-NOV-1999; 99WO-US028214.
PR 30-NOV-1999; 99WO-US028313.
PR 01-DEC-1999; 99WO-US028301.
PR 02-DEC-1999; 99WO-US028564.
PR 02-DEC-1999; 99WO-US028565.
PR 16-DEC-1999; 99WO-US030095.
PR 20-DEC-1999; 99WO-US030911.
PR 20-DEC-1999; 99WO-US030999.
PR 05-JAN-2000; 2000WO-US000219.
PR 11-FEB-2000; 2000WO-US003565.
PR 22-FEB-2000; 2000WO-US004414.
PR 24-FEB-2000; 2000WO-US005004.
PR 02-MAR-2000; 2000WO-US005841.
PR 20-MAR-2000; 2000WO-US007377.
PR 30-MAR-2000; 2000WO-US008439.
PR 22-MAY-2000; 2000WO-US014042.
PR 02-JUN-2000; 2000WO-US015264.
PR 28-JUL-2000; 2000WO-US020710.
PR 24-AUG-2000; 2000WO-US023328.
PR 18-SEP-2000; 2000US-00665350.
XX
PA (GETH) GENENTECH INC.
XX
PI Ashkenazi A, Botstein D, Desnoyers L, Eaton DL, Ferrara N;
PI Filvaroff E, Fong S, Gao W, Gerber H, Gerritsen ME, Goddard A;
PI Godowski PJ, Grimaldi JC, Gurney AL, Hillan KJ, Kljavin IJ;
PI Mather JP, Pan J, Paoni NF, Roy MA, Stewart TA, Tumas D;
PI Williams PM, Wood WI;
XX
DR WPI; 2003-370793/35.
DR N-PSDB; ACD07512.
XX
PT New genes and secreted and transmembrane polypeptides (e.g. PRO245 or
PT PRO335), useful for treating or diagnosing e.g. Alzheimer's disease,
PT cancers, hemorrhage, rheumatoid arthritis, diabetes, cirrhosis, ischemia
PT or strokes.
XX
PS Claim 12; Fig 50; 482pp; English.
XX
CC The invention describes a new isolated nucleic acid molecule comprising
CC the full length coding sequence of the DNA deposited with the American
CC Type Culture Collection (e.g. ATCC Deposit No. 209258) ,or a sequence
CC with at least 80% identity to a DNA encoding a PRO polypeptide comprising
CC any of 61 sequences having 164-1119 amino acids fully defined in the
CC specification. The PRO polypeptides or polynucleotides are useful as
CC pharmaceuticals, diagnostics, biosensors or bioreactors. These are
CC particularly useful for detecting or treating e.g. Parkinson's disease,
CC Alzheimer's disease, inflammations, nephritis, wound healing, nerve
CC repair, collateral blood vessel formation, cancers (e.g. colorectal
CC cancer), haemorrhage (or reduce risk for haemorrhage), rheumatoid
CC arthritis, diabetes, cirrhosis of the liver, fibrosis of the lungs,
CC restenosis, dermal fibrotic conditions (e.g. keloids or scarring),
CC ischaemia, strokes, hypertension, heart attacks, atherosclerosis, or
CC infertility in mammals (e.g. humans, dogs, cats, cattle, horses, sheep,
CC pigs, goats, or rabbits) The PRO polypeptides are useful as targets for
CC therapeutic intervention in these diseases, and diagnostic determination
CC of the presence of these diseases. The PRO polypeptides are also useful
CC as molecular weight markers, or for chromosome identification. The PRO
CC genes are useful as hybridisation probes, or for screening libraries of
CC human cDNA, genomic DNA or mRNA. The PRO genes may also be used in gene
CC therapy, particularly for replacing a defective gene. This is the amino
CC acid sequence of a novel human secreted and transmembrane PRO polypeptide
XX
SQ Sequence 382 AA;

Query Match 98.7%; Score 1101; DB 6; Length 382;
Best Local Similarity 96.3%; Pred. No. 5.2e-103;
Matches 206; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

```
Qy          1 ATGRLLS--------GQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 52
              |||||||        |||||||||||||||||||||||||||||||||||||||||||||
Db         22 ATGRLLSASDLDLRGGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 81

Qy         53 ESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVD 112
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 ESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVD 141

Qy        113 EPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGE 172
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 EPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGE 201

Qy        173 ETELTTPVLPEETQEEDAKKTFKESREAALNLAY 206
              |||||||||||||||||||||||||||||||||
Db        202 ETELTTPVLPEETQEEDAKKTFKESREAALNLAY 235
```

```
RESULT 40
ABR99361
ID   ABR99361 standard; protein; 382 AA.
XX
AC   ABR99361;
XX
DT   18-SEP-2003  (first entry)
XX
DE   Human secreted polypeptide PRO234, SEQ ID NO:20.
XX
KW   KW Human; PRO; secreted protein; transmembrane protein;
KW   extracellular domain; tumour necrosis factor-alpha; TNF-alpha;
KW   chondrocyte; proliferation; differentiation; cartilage disorder;
KW   bone disorder; arthritis; sports injury; cancer; tumour; diagnosis;
KW   adrenal tumour; lung; colon; breast; prostate; kidney; rectum; cervix;
KW   liver; drug screening; transgenic animal; genetic analysis;
KW   antiarthritic; vulnerary; gene therapy.
XX
OS   Homo sapiens.
XX
PN   US2003040063-A1.
XX
PD   27-FEB-2003.
XX
PF   26-JUN-2002; 2002US-00183006.
XX
PR   18-SEP-1997;   97US-0059263P.
PR   18-SEP-1997;   97US-0059266P.
PR   17-OCT-1997;   97US-0062250P.
PR   21-OCT-1997;   97US-0063486P.
PR   24-OCT-1997;   97US-0063120P.
PR   24-OCT-1997;   97US-0063121P.
PR   28-OCT-1997;   97US-0063540P.
PR   28-OCT-1997;   97US-0063541P.
PR   28-OCT-1997;   97US-0063544P.
PR   28-OCT-1997;   97US-0063564P.
PR   29-OCT-1997;   97US-0063734P.
PR   31-OCT-1997;   97US-0063870P.
PR   31-OCT-1997;   97US-0064103P.
PR   13-NOV-1997;   97US-0065311P.
PR   21-NOV-1997;   97US-0066120P.
PR   24-NOV-1997;   97US-0066466P.
PR   24-NOV-1997;   97US-0066772P.
PR   11-DEC-1997;   97US-0069335P.
PR   12-DEC-1997;   97US-0069425P.
PR   17-DEC-1997;   97US-0069870P.
PR   18-DEC-1997;   97US-0068017P.
PR   10-MAR-1998;   98US-0077450P.
PR   11-MAR-1998;   98US-0077632P.
PR   11-MAR-1998;   98US-0077649P.
PR   20-MAR-1998;   98US-0078886P.
PR   20-MAR-1998;   98US-0078939P.
PR   27-MAR-1998;   98US-0079664P.
PR   27-MAR-1998;   98US-0079786P.
PR   31-MAR-1998;   98US-0080107P.
PR   31-MAR-1998;   98US-0080194P.
PR   01-APR-1998;   98US-0080327P.
PR   01-APR-1998;   98US-0080333P.
PR   08-APR-1998;   98US-0081049P.
PR   08-APR-1998;   98US-0081070P.
PR   09-APR-1998;   98US-0081195P.
PR   15-APR-1998;   98US-0081838P.
PR   21-APR-1998;   98US-0082568P.
PR   21-APR-1998;   98US-0082569P.
PR   22-APR-1998;   98US-0082704P.
PR   22-APR-1998;   98US-0082797P.
PR   28-APR-1998;   98US-0083322P.
PR   29-APR-1998;   98US-0083495P.
PR   29-APR-1998;   98US-0083496P.
PR   29-APR-1998;   98US-0083499P.
PR   29-APR-1998;   98US-0083559P.
PR   05-MAY-1998;   98US-0084366P.
PR   06-MAY-1998;   98US-0084414P.
PR   07-MAY-1998;   98US-0084639P.
PR   07-MAY-1998;   98US-0084640P.
PR   07-MAY-1998;   98US-0084643P.
PR   15-MAY-1998;   98US-0085579P.
PR   15-MAY-1998;   98US-0085580P.
PR   15-MAY-1998;   98US-0085582P.
PR   15-MAY-1998;   98US-0085700P.
PR   18-MAY-1998;   98US-0086023P.
PR   22-MAY-1998;   98US-0086392P.
PR   22-MAY-1998;   98US-0086486P.
PR   28-MAY-1998;   98US-0087098P.
PR   28-MAY-1998;   98US-0087208P.
PR   02-JUN-1998;   98US-0087609P.
PR   02-JUN-1998;   98US-0087759P.
PR   03-JUN-1998;   98US-0087827P.
PR   04-JUN-1998;   98US-0088025P.
PR   04-JUN-1998;   98US-0088028P.
PR   04-JUN-1998;   98US-0088029P.
PR   04-JUN-1998;   98US-0088033P.
PR   04-JUN-1998;   98US-0088326P.
PR   05-JUN-1998;   98US-0088167P.
PR   05-JUN-1998;   98US-0088202P.
PR   05-JUN-1998;   98US-0088212P.
PR   05-JUN-1998;   98US-0088217P.
PR   09-JUN-1998;   98US-0088655P.
PR   10-JUN-1998;   98US-0088722P.
PR   10-JUN-1998;   98US-0088738P.
PR   10-JUN-1998;   98US-0088740P.
PR   10-JUN-1998;   98US-0088811P.
PR   10-JUN-1998;   98US-0088824P.
PR   10-JUN-1998;   98US-0088825P.
PR   10-JUN-1998;   98US-0088826P.
PR   11-JUN-1998;   98US-0088861P.
PR   11-JUN-1998;   98US-0088863P.
PR   11-JUN-1998;   98US-0088876P.
PR   12-JUN-1998;   98US-0089090P.
PR   12-JUN-1998;   98US-0089105P.
PR   16-JUN-1998;   98US-0089512P.
PR   16-JUN-1998;   98US-0089514P.
PR   17-JUN-1998;   98US-0089538P.
PR   17-JUN-1998;   98US-0089598P.
PR   17-JUN-1998;   98US-0089653P.
PR   18-JUN-1998;   98US-0089908P.
PR   19-JUN-1998;   98US-0089952P.
PR   22-JUN-1998;   98US-0090246P.
PR   22-JUN-1998;   98US-0090252P.
PR   22-JUN-1998;   98US-0090254P.
PR   24-JUN-1998;   98US-0090429P.
PR   24-JUN-1998;   98US-0090435P.
PR   24-JUN-1998;   98US-0090444P.
PR   24-JUN-1998;   98US-0090461P.
PR   24-JUN-1998;   98US-0090535P.
PR   24-JUN-1998;   98US-0090540P.
PR   25-JUN-1998;   98US-0090676P.
PR   25-JUN-1998;   98US-0090678P.
PR   25-JUN-1998;   98US-0090688P.
PR   25-JUN-1998;   98US-0090690P.
PR   25-JUN-1998;   98US-0090694P.
PR   25-JUN-1998;   98US-0090695P.
PR   25-JUN-1998;   98US-0090696P.
PR   26-JUN-1998;   98US-00105413.
PR   26-JUN-1998;   98US-0090862P.
PR   26-JUN-1998;   98US-0090863P.
PR   26-JUN-1998;   98US-0091010P.
PR   01-JUL-1998;   98US-0091359P.
PR   01-JUL-1998;   98US-0091544P.
PR   02-JUL-1998;   98US-0091478P.
PR   02-JUL-1998;   98US-0091486P.
PR   02-JUL-1998;   98US-0091626P.
PR   02-JUL-1998;   98US-0091628P.
PR   02-JUL-1998;   98US-0091632P.
PR   24-JUL-1998;   98US-0094006P.
```

PR 04-AUG-1998; 98US-0095282P.
PR 10-AUG-1998; 98US-0095998P.
PR 10-AUG-1998; 98US-0096012P.
PR 17-AUG-1998; 98US-0096757P.
PR 17-AUG-1998; 98US-0096766P.
PR 17-AUG-1998; 98US-0096867P.
PR 17-AUG-1998; 98US-0096891P.
PR 17-AUG-1998; 98US-0096897P.
PR 18-AUG-1998; 98US-0096949P.
PR 18-AUG-1998; 98US-0096959P.
PR 18-AUG-1998; 98US-0097022P.
PR 26-AUG-1998; 98US-0097952P.
PR 26-AUG-1998; 98US-0097954P.
PR 26-AUG-1998; 98US-0097955P.
PR 26-AUG-1998; 98US-0097971P.
PR 26-AUG-1998; 98US-0097974P.
PR 26-AUG-1998; 98US-0098014P.
PR 01-SEP-1998; 98US-0098716P.
PR 01-SEP-1998; 98US-0098723P.
PR 02-SEP-1998; 98US-0098803P.
PR 02-SEP-1998; 98US-0098821P.
PR 02-SEP-1998; 98US-0098843P.
PR 09-SEP-1998; 98US-0099602P.
PR 10-SEP-1998; 98US-0099741P.
PR 10-SEP-1998; 98US-0099754P.
PR 10-SEP-1998; 98US-0099763P.
PR 10-SEP-1998; 98US-0099812P.
PR 15-SEP-1998; 98US-0100388P.
PR 16-SEP-1998; 98US-0100662P.
PR 16-SEP-1998; 98US-0100664P.
PR 16-SEP-1998; 98US-0101751P.
PR 16-SEP-1998; 98WO-US019330.
PR 17-SEP-1998; 98US-0100683P.
PR 17-SEP-1998; 98US-0100684P.
PR 17-SEP-1998; 98US-0100919P.
PR 17-SEP-1998; 98US-0100930P.
PR 18-SEP-1998; 98US-0100849P.
PR 18-SEP-1998; 98US-0101014P.
PR 18-SEP-1998; 98US-0101068P.
PR 23-SEP-1998; 98US-0101471P.
PR 23-SEP-1998; 98US-0101472P.
PR 23-SEP-1998; 98US-0101475P.
PR 23-SEP-1998; 98US-0101477P.
PR 24-SEP-1998; 98US-0101738P.
PR 24-SEP-1998; 98US-0101739P.
PR 24-SEP-1998; 98US-0101743P.
PR 24-SEP-1998; 98US-0101922P.
PR 25-SEP-1998; 98US-0101786P.
PR 29-SEP-1998; 98US-0102207P.
PR 29-SEP-1998; 98US-0102240P.
PR 29-SEP-1998; 98US-0102330P.
PR 29-SEP-1998; 98US-0102331P.
PR 30-SEP-1998; 98US-0102487P.
PR 30-SEP-1998; 98US-0102570P.
PR 30-SEP-1998; 98US-0102571P.
PR 01-OCT-1998; 98US-0102684P.
PR 01-OCT-1998; 98US-0102687P.

Query Match 98.7%; Score 1101; DB 6; Length 382;
Best Local Similarity 96.3%; Pred. No. 5.2e-103;
Matches 206; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

```
Qy          1 ATGRLLS--------GQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 52
              |||||||        ||||||||||||||||||||||||||||||||||||||||||||
Db         22 ATGRLLSASDLDLRGGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 81

Qy         53 ESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVD 112
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 ESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVD 141

Qy        113 EPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGE 172
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 EPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGE 201
```

Qy 173 ETELTTPVLPEETQEEDAKKTFKESREAALNLAY 206
||||||||||||||||||||||||||||||||||
Db 202 ETELTTPVLPEETQEEDAKKTFKESREAALNLAY 235

RESULT 41
ABR98751
ID ABR98751 standard; protein; 382 AA.
XX
AC ABR98751;
XX
DT 17-SEP-2003 (first entry)
XX
DE Human secreted polypeptide PRO234, SEQ ID NO:20.
XX
KW Human; PRO; secreted protein; transmembrane protein;
KW extracellular domain; tumour necrosis factor-alpha; TNF-alpha;
KW chondrocyte; proliferation; differentiation; cartilage disorder;
KW bone disorder; arthritis; sports injury; cancer; tumour; diagnosis;
KW adrenal tumour; lung; colon; breast; prostate; kidney; rectum; cervix;
KW liver; drug screening; transgenic animal; genetic analysis;
KW antiarthritic; vulnerary; gene therapy.
XX
OS Homo sapiens.
XX
PN US2003040064-A1.
XX
PD 27-FEB-2003.
XX
PF 26-JUN-2002; 2002US-00183008.
XX
PR 18-SEP-1997; 97US-0059263P.
PR 18-SEP-1997; 97US-0059266P.
PR 17-OCT-1997; 97US-0062250P.
PR 21-OCT-1997; 97US-0063486P.
PR 24-OCT-1997; 97US-0063120P.
PR 24-OCT-1997; 97US-0063121P.
PR 28-OCT-1997; 97US-0063540P.
PR 28-OCT-1997; 97US-0063541P.
PR 28-OCT-1997; 97US-0063544P.
PR 28-OCT-1997; 97US-0063564P.
PR 29-OCT-1997; 97US-0063734P.
PR 31-OCT-1997; 97US-0063870P.
PR 31-OCT-1997; 97US-0064103P.
PR 13-NOV-1997; 97US-0065311P.
PR 21-NOV-1997; 97US-0066120P.
PR 24-NOV-1997; 97US-0066466P.
PR 24-NOV-1997; 97US-0066772P.
PR 11-DEC-1997; 97US-0069335P.
PR 12-DEC-1997; 97US-0069425P.
PR 17-DEC-1997; 97US-0069870P.
PR 18-DEC-1997; 97US-0068017P.
PR 10-MAR-1998; 98US-0077450P.
PR 11-MAR-1998; 98US-0077632P.
PR 11-MAR-1998; 98US-0077649P.
PR 20-MAR-1998; 98US-0078886P.
PR 20-MAR-1998; 98US-0078939P.
PR 27-MAR-1998; 98US-0079664P.
PR 27-MAR-1998; 98US-0079786P.
PR 31-MAR-1998; 98US-0080107P.
PR 31-MAR-1998; 98US-0080194P.
PR 01-APR-1998; 98US-0080327P.
PR 01-APR-1998; 98US-0080333P.
PR 08-APR-1998; 98US-0081049P.
PR 08-APR-1998; 98US-0081070P.
PR 09-APR-1998; 98US-0081195P.
PR 15-APR-1998; 98US-0081838P.
PR 21-APR-1998; 98US-0082568P.
PR 21-APR-1998; 98US-0082569P.
PR 22-APR-1998; 98US-0082704P.
PR 22-APR-1998; 98US-0082797P.
PR 28-APR-1998; 98US-0083322P.

PR 29-APR-1998; 98US-0083495P.
PR 29-APR-1998; 98US-0083496P.
PR 29-APR-1998; 98US-0083499P.
PR 29-APR-1998; 98US-0083559P.
PR 05-MAY-1998; 98US-0084366P.
PR 06-MAY-1998; 98US-0084414P.
PR 07-MAY-1998; 98US-0084639P.
PR 07-MAY-1998; 98US-0084640P.
PR 07-MAY-1998; 98US-0084643P.
PR 15-MAY-1998; 98US-0085579P.
PR 15-MAY-1998; 98US-0085580P.
PR 15-MAY-1998; 98US-0085582P.
PR 15-MAY-1998; 98US-0085700P.
PR 18-MAY-1998; 98US-0086023P.
PR 22-MAY-1998; 98US-0086392P.
PR 22-MAY-1998; 98US-0086486P.
PR 28-MAY-1998; 98US-0087098P.
PR 28-MAY-1998; 98US-0087208P.
PR 02-JUN-1998; 98US-0087609P.
PR 02-JUN-1998; 98US-0087759P.
PR 03-JUN-1998; 98US-0087827P.
PR 04-JUN-1998; 98US-0088025P.
PR 04-JUN-1998; 98US-0088028P.
PR 04-JUN-1998; 98US-0088029P.
PR 04-JUN-1998; 98US-0088033P.
PR 04-JUN-1998; 98US-0088326P.
PR 05-JUN-1998; 98US-0088167P.
PR 05-JUN-1998; 98US-0088202P.
PR 05-JUN-1998; 98US-0088212P.
PR 05-JUN-1998; 98US-0088217P.
PR 09-JUN-1998; 98US-0088655P.
PR 10-JUN-1998; 98US-0088722P.
PR 10-JUN-1998; 98US-0088738P.
PR 10-JUN-1998; 98US-0088740P.
PR 10-JUN-1998; 98US-0088811P.
PR 10-JUN-1998; 98US-0088824P.
PR 10-JUN-1998; 98US-0088825P.
PR 10-JUN-1998; 98US-0088826P.
PR 11-JUN-1998; 98US-0088861P.
PR 11-JUN-1998; 98US-0088863P.
PR 11-JUN-1998; 98US-0088876P.
PR 12-JUN-1998; 98US-0089090P.
PR 12-JUN-1998; 98US-0089105P.
PR 16-JUN-1998; 98US-0089512P.
PR 16-JUN-1998; 98US-0089514P.
PR 17-JUN-1998; 98US-0089538P.
PR 17-JUN-1998; 98US-0089598P.
PR 17-JUN-1998; 98US-0089653P.
PR 18-JUN-1998; 98US-0089908P.
PR 19-JUN-1998; 98US-0089952P.
PR 22-JUN-1998; 98US-0090246P.
PR 22-JUN-1998; 98US-0090252P.
PR 22-JUN-1998; 98US-0090254P.
PR 24-JUN-1998; 98US-0090429P.
PR 24-JUN-1998; 98US-0090435P.
PR 24-JUN-1998; 98US-0090444P.
PR 24-JUN-1998; 98US-0090461P.
PR 24-JUN-1998; 98US-0090535P.
PR 24-JUN-1998; 98US-0090540P.
PR 25-JUN-1998; 98US-0090676P.
PR 25-JUN-1998; 98US-0090678P.
PR 25-JUN-1998; 98US-0090688P.
PR 25-JUN-1998; 98US-0090690P.
PR 25-JUN-1998; 98US-0090694P.
PR 25-JUN-1998; 98US-0090695P.
PR 25-JUN-1998; 98US-0090696P.
PR 26-JUN-1998; 98US-00105413.
PR 26-JUN-1998; 98US-0090862P.
PR 26-JUN-1998; 98US-0090863P.
PR 26-JUN-1998; 98US-0091010P.
PR 01-JUL-1998; 98US-0091359P.
PR 01-JUL-1998; 98US-0091544P.
PR 02-JUL-1998; 98US-0091478P.
PR 02-JUL-1998; 98US-0091486P.
PR 02-JUL-1998; 98US-0091626P.
PR 02-JUL-1998; 98US-0091628P.
PR 02-JUL-1998; 98US-0091632P.
PR 24-JUL-1998; 98US-0094006P.
PR 04-AUG-1998; 98US-0095282P.
PR 10-AUG-1998; 98US-0095998P.
PR 10-AUG-1998; 98US-0096012P.
PR 17-AUG-1998; 98US-0096757P.
PR 17-AUG-1998; 98US-0096766P.
PR 17-AUG-1998; 98US-0096867P.
PR 17-AUG-1998; 98US-0096891P.
PR 17-AUG-1998; 98US-0096897P.
PR 18-AUG-1998; 98US-0096949P.
PR 18-AUG-1998; 98US-0096959P.
PR 18-AUG-1998; 98US-0097022P.
PR 26-AUG-1998; 98US-0097952P.
PR 26-AUG-1998; 98US-0097954P.
PR 26-AUG-1998; 98US-0097955P.
PR 26-AUG-1998; 98US-0097971P.
PR 26-AUG-1998; 98US-0097974P.
PR 26-AUG-1998; 98US-0098014P.
PR 01-SEP-1998; 98US-0098716P.
PR 01-SEP-1998; 98US-0098723P.
PR 02-SEP-1998; 98US-0098803P.
PR 02-SEP-1998; 98US-0098821P.
PR 02-SEP-1998; 98US-0098843P.
PR 09-SEP-1998; 98US-0099602P.
PR 10-SEP-1998; 98US-0099741P.
PR 10-SEP-1998; 98US-0099754P.
PR 10-SEP-1998; 98US-0099763P.
PR 10-SEP-1998; 98US-0099812P.
PR 15-SEP-1998; 98US-0100388P.
PR 16-SEP-1998; 98US-0100662P.
PR 16-SEP-1998; 98US-0100664P.
PR 16-SEP-1998; 98US-0101751P.
PR 16-SEP-1998; 98WO-US019330.
PR 17-SEP-1998; 98US-0100683P.
PR 17-SEP-1998; 98US-0100684P.
PR 17-SEP-1998; 98US-0100919P.
PR 17-SEP-1998; 98US-0100930P.
PR 18-SEP-1998; 98US-0100849P.
PR 18-SEP-1998; 98US-0101014P.
PR 18-SEP-1998; 98US-0101068P.
PR 23-SEP-1998; 98US-0101471P.
PR 23-SEP-1998; 98US-0101472P.
PR 23-SEP-1998; 98US-0101475P.
PR 23-SEP-1998; 98US-0101477P.
PR 24-SEP-1998; 98US-0101738P.
PR 24-SEP-1998; 98US-0101739P.
PR 24-SEP-1998; 98US-0101743P.
PR 24-SEP-1998; 98US-0101922P.
PR 25-SEP-1998; 98US-0101786P.
PR 29-SEP-1998; 98US-0102207P.
PR 29-SEP-1998; 98US-0102240P.
PR 29-SEP-1998; 98US-0102330P.
PR 29-SEP-1998; 98US-0102331P.
PR 30-SEP-1998; 98US-0102487P.
PR 30-SEP-1998; 98US-0102570P.
PR 30-SEP-1998; 98US-0102571P.
PR 01-OCT-1998; 98US-0102684P.
PR 01-OCT-1998; 98US-0102687P.

Query Match 98.7%; Score 1101; DB 6; Length 382;
Best Local Similarity 96.3%; Pred. No. 5.2e-103;
Matches 206; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

```
Qy          1 ATGRLLS--------GQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 52
              |||||||        ||||||||||||||||||||||||||||||||||||||||||||
Db         22 ATGRLLSASDLDLRGGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 81

Qy         53 ESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVD 112
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        82 ESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVD 141

Qy       113 EPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGE 172
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       142 EPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGE 201

Qy       173 ETELTTPVLPEETQEEDAKKTFKESREAALNLAY 206
            |||||||||||||||||||||||||||||||||
Db       202 ETELTTPVLPEETQEEDAKKTFKESREAALNLAY 235


RESULT 42
ABO16274
ID   ABO16274 standard; protein; 382 AA.
XX
AC   ABO16274;
XX
DT   25-AUG-2003  (first entry)
XX
DE   Human secreted/transmembrane protein (PRO) #10.
XX
KW   Human; secreted and transmembrane protein; PRO; TNF-alpha;
KW   tumour necrosis factor alpha; chondrocyte cell; tumour; gene therapy;
KW   tissue typing; adrenal tumour; lung tumour; colon tumour; breast tumour;
KW   prostate tumour; rectal tumour; cervical tumour; liver tumour.
XX
OS   Homo sapiens.
XX
PN   US2003027267-A1.
XX
PD   06-FEB-2003.
XX
PF   19-JUN-2002; 2002US-00175739.
XX
PR   18-SEP-1997;   97US-0059263P.
PR   18-SEP-1997;   97US-0059266P.
PR   17-OCT-1997;   97US-0062250P.
PR   21-OCT-1997;   97US-0063486P.
PR   24-OCT-1997;   97US-0063120P.
PR   24-OCT-1997;   97US-0063121P.
PR   28-OCT-1997;   97US-0063540P.
PR   28-OCT-1997;   97US-0063541P.
PR   28-OCT-1997;   97US-0063544P.
PR   28-OCT-1997;   97US-0063564P.
PR   29-OCT-1997;   97US-0063734P.
PR   31-OCT-1997;   97US-0063870P.
PR   31-OCT-1997;   97US-0064103P.
PR   13-NOV-1997;   97US-0065311P.
PR   21-NOV-1997;   97US-0066120P.
PR   24-NOV-1997;   97US-0066466P.
PR   24-NOV-1997;   97US-0066772P.
PR   11-DEC-1997;   97US-0069335P.
PR   12-DEC-1997;   97US-0069425P.
PR   17-DEC-1997;   97US-0069870P.
PR   18-DEC-1997;   97US-0068017P.
PR   10-MAR-1998;   98US-0077450P.
PR   11-MAR-1998;   98US-0077632P.
PR   11-MAR-1998;   98US-0077649P.
PR   20-MAR-1998;   98US-0078886P.
PR   20-MAR-1998;   98US-0078939P.
PR   27-MAR-1998;   98US-0079664P.
PR   27-MAR-1998;   98US-0079786P.
PR   31-MAR-1998;   98US-0080107P.
PR   31-MAR-1998;   98US-0080194P.
PR   01-APR-1998;   98US-0080327P.
PR   01-APR-1998;   98US-0080333P.
PR   08-APR-1998;   98US-0081049P.
PR   08-APR-1998;   98US-0081070P.
PR   09-APR-1998;   98US-0081195P.
PR   15-APR-1998;   98US-0081838P.
PR   21-APR-1998;   98US-0082568P.
PR   21-APR-1998;   98US-0082569P.
PR   22-APR-1998;   98US-0082704P.
PR   22-APR-1998;   98US-0082797P.
PR   28-APR-1998;   98US-0083322P.
PR   29-APR-1998;   98US-0083495P.
PR   29-APR-1998;   98US-0083496P.
PR   29-APR-1998;   98US-0083499P.
PR   29-APR-1998;   98US-0083559P.
PR   05-MAY-1998;   98US-0084366P.
PR   06-MAY-1998;   98US-0084414P.
PR   07-MAY-1998;   98US-0084639P.
PR   07-MAY-1998;   98US-0084640P.
PR   07-MAY-1998;   98US-0084643P.
PR   15-MAY-1998;   98US-0085579P.
PR   15-MAY-1998;   98US-0085580P.
PR   15-MAY-1998;   98US-0085582P.
PR   15-MAY-1998;   98US-0085700P.
PR   18-MAY-1998;   98US-0086023P.
PR   22-MAY-1998;   98US-0086392P.
PR   22-MAY-1998;   98US-0086486P.
PR   28-MAY-1998;   98US-0087098P.
PR   28-MAY-1998;   98US-0087208P.
PR   02-JUN-1998;   98US-0087609P.
PR   02-JUN-1998;   98US-0087759P.
PR   03-JUN-1998;   98US-0087827P.
PR   04-JUN-1998;   98US-0088025P.
PR   04-JUN-1998;   98US-0088028P.
PR   04-JUN-1998;   98US-0088029P.
PR   04-JUN-1998;   98US-0088033P.
PR   04-JUN-1998;   98US-0088326P.
PR   05-JUN-1998;   98US-0088167P.
PR   05-JUN-1998;   98US-0088202P.
PR   05-JUN-1998;   98US-0088212P.
PR   05-JUN-1998;   98US-0088217P.
PR   09-JUN-1998;   98US-0088655P.
PR   10-JUN-1998;   98US-0088722P.
PR   10-JUN-1998;   98US-0088738P.
PR   10-JUN-1998;   98US-0088740P.
PR   10-JUN-1998;   98US-0088811P.
PR   10-JUN-1998;   98US-0088824P.
PR   10-JUN-1998;   98US-0088825P.
PR   10-JUN-1998;   98US-0088826P.
PR   11-JUN-1998;   98US-0088861P.
PR   11-JUN-1998;   98US-0088863P.
PR   11-JUN-1998;   98US-0088876P.
PR   12-JUN-1998;   98US-0089090P.
PR   12-JUN-1998;   98US-0089105P.
PR   16-JUN-1998;   98US-0089512P.
PR   16-JUN-1998;   98US-0089514P.
PR   17-JUN-1998;   98US-0089538P.
PR   17-JUN-1998;   98US-0089598P.
PR   17-JUN-1998;   98US-0089653P.
PR   18-JUN-1998;   98US-0089908P.
PR   19-JUN-1998;   98US-0089952P.
PR   22-JUN-1998;   98US-0090246P.
PR   22-JUN-1998;   98US-0090252P.
PR   22-JUN-1998;   98US-0090254P.
PR   24-JUN-1998;   98US-0090429P.
PR   24-JUN-1998;   98US-0090435P.
PR   24-JUN-1998;   98US-0090444P.
PR   24-JUN-1998;   98US-0090461P.
PR   24-JUN-1998;   98US-0090535P.
PR   24-JUN-1998;   98US-0090540P.
PR   25-JUN-1998;   98US-0090676P.
PR   25-JUN-1998;   98US-0090678P.
PR   25-JUN-1998;   98US-0090688P.
PR   25-JUN-1998;   98US-0090690P.
PR   25-JUN-1998;   98US-0090694P.
PR   25-JUN-1998;   98US-0090695P.
PR   25-JUN-1998;   98US-0090696P.
PR   26-JUN-1998;   98US-00105413.
PR   26-JUN-1998;   98US-0090862P.
PR   26-JUN-1998;   98US-0090863P.
PR   26-JUN-1998;   98US-0091010P.
PR   01-JUL-1998;   98US-0091359P.
```

PR 01-JUL-1998; 98US-0091544P.
PR 02-JUL-1998; 98US-0091478P.
PR 02-JUL-1998; 98US-0091486P.
PR 02-JUL-1998; 98US-0091626P.
PR 02-JUL-1998; 98US-0091628P.
PR 02-JUL-1998; 98US-0091632P.
PR 24-JUL-1998; 98US-0094006P.
PR 04-AUG-1998; 98US-0095282P.
PR 10-AUG-1998; 98US-0095998P.
PR 10-AUG-1998; 98US-0096012P.
PR 17-AUG-1998; 98US-0096757P.
PR 17-AUG-1998; 98US-0096766P.
PR 17-AUG-1998; 98US-0096867P.
PR 17-AUG-1998; 98US-0096891P.
PR 17-AUG-1998; 98US-0096897P.
PR 18-AUG-1998; 98US-0096949P.
PR 18-AUG-1998; 98US-0096959P.
PR 18-AUG-1998; 98US-0097022P.
PR 26-AUG-1998; 98US-0097952P.
PR 26-AUG-1998; 98US-0097954P.
PR 26-AUG-1998; 98US-0097955P.
PR 26-AUG-1998; 98US-0097971P.
PR 26-AUG-1998; 98US-0097974P.
PR 26-AUG-1998; 98US-0098014P.
PR 01-SEP-1998; 98US-0098716P.
PR 01-SEP-1998; 98US-0098723P.
PR 02-SEP-1998; 98US-0098803P.
PR 02-SEP-1998; 98US-0098821P.
PR 02-SEP-1998; 98US-0098843P.
PR 09-SEP-1998; 98US-0099602P.
PR 10-SEP-1998; 98US-0099741P.
PR 10-SEP-1998; 98US-0099754P.
PR 10-SEP-1998; 98US-0099763P.
PR 10-SEP-1998; 98US-0099812P.
PR 15-SEP-1998; 98US-0100388P.
PR 16-SEP-1998; 98US-0100662P.
PR 16-SEP-1998; 98US-0100664P.
PR 16-SEP-1998; 98US-0101751P.
PR 16-SEP-1998; 98WO-US019330.
PR 17-SEP-1998; 98US-0100683P.
PR 17-SEP-1998; 98US-0100684P.
PR 17-SEP-1998; 98US-0100919P.
PR 17-SEP-1998; 98US-0100930P.
PR 18-SEP-1998; 98US-0100849P.
PR 18-SEP-1998; 98US-0101014P.
PR 18-SEP-1998; 98US-0101068P.
PR 23-SEP-1998; 98US-0101471P.
PR 23-SEP-1998; 98US-0101472P.
PR 23-SEP-1998; 98US-0101475P.
PR 23-SEP-1998; 98US-0101477P.
PR 24-SEP-1998; 98US-0101738P.
PR 24-SEP-1998; 98US-0101739P.
PR 24-SEP-1998; 98US-0101743P.
PR 24-SEP-1998; 98US-0101922P.
PR 25-SEP-1998; 98US-0101786P.
PR 29-SEP-1998; 98US-0102207P.
PR 29-SEP-1998; 98US-0102240P.
PR 29-SEP-1998; 98US-0102330P.
PR 29-SEP-1998; 98US-0102331P.
PR 30-SEP-1998; 98US-0102487P.
PR 30-SEP-1998; 98US-0102570P.
PR 30-SEP-1998; 98US-0102571P.
PR 01-OCT-1998; 98US-0102684P.
PR 01-OCT-1998; 98US-0102687P.
PR 02-OCT-1998; 98US-0102965P.
PR 06-OCT-1998; 98US-0103258P.
PR 06-OCT-1998; 98US-0103449P.

Query Match 98.7%; Score 1101; DB 6; Length 382;
Best Local Similarity 96.3%; Pred. No. 5.2e-103;
Matches 206; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

```
Qy          1 ATGRLLS--------GQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 52
              |||||||        ||||||||||||||||||||||||||||||||||||||||||||
Db         22 ATGRLLSASDLDLRGGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 81

Qy         53 ESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVD 112
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 ESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVD 141

Qy        113 EPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGE 172
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 EPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGE 201

Qy        173 ETELTTPVLPEETQEEDAKKTFKESREAALNLAY 206
              |||||||||||||||||||||||||||||||||
Db        202 ETELTTPVLPEETQEEDAKKTFKESREAALNLAY 235
```

RESULT 43
ABR92174
ID ABR92174 standard; protein; 382 AA.
XX
AC ABR92174;
XX
DT 13-SEP-2003 (first entry)
XX
DE Human secreted polypeptide PRO234, SEQ ID NO:20.
XX
KW Human; PRO; secreted protein; transmembrane protein;
KW extracellular domain; tumour necrosis factor-alpha; TNF-alpha;
KW chondrocyte; proliferation; differentiation; cartilage disorder;
KW bone disorder; arthritis; sports injury; cancer; tumour; diagnosis;
KW adrenal tumour; lung; colon; breast; prostate; kidney; rectum; cervix;
KW liver; drug screening; transgenic animal; genetic analysis;
KW antiarthritic; vulnerary; gene therapy.
XX
OS Homo sapiens.
XX
PN US2003036160-A1.
XX
PD 20-FEB-2003.
XX
PF 02-JUL-2002; 2002US-00188781.
XX
PR 18-SEP-1997; 97US-0059263P.
PR 18-SEP-1997; 97US-0059266P.
PR 17-OCT-1997; 97US-0062250P.
PR 21-OCT-1997; 97US-0063486P.
PR 24-OCT-1997; 97US-0063120P.
PR 24-OCT-1997; 97US-0063121P.
PR 28-OCT-1997; 97US-0063540P.
PR 28-OCT-1997; 97US-0063541P.
PR 28-OCT-1997; 97US-0063544P.
PR 28-OCT-1997; 97US-0063564P.
PR 29-OCT-1997; 97US-0063734P.
PR 31-OCT-1997; 97US-0063870P.
PR 31-OCT-1997; 97US-0064103P.
PR 13-NOV-1997; 97US-0065311P.
PR 21-NOV-1997; 97US-0066120P.
PR 24-NOV-1997; 97US-0066466P.
PR 24-NOV-1997; 97US-0066772P.
PR 11-DEC-1997; 97US-0069335P.
PR 12-DEC-1997; 97US-0069425P.
PR 17-DEC-1997; 97US-0069870P.
PR 18-DEC-1997; 97US-0068017P.
PR 10-MAR-1998; 98US-0077450P.
PR 11-MAR-1998; 98US-0077632P.
PR 11-MAR-1998; 98US-0077649P.
PR 20-MAR-1998; 98US-0078886P.
PR 20-MAR-1998; 98US-0078939P.
PR 27-MAR-1998; 98US-0079664P.
PR 27-MAR-1998; 98US-0079786P.
PR 31-MAR-1998; 98US-0080107P.
PR 31-MAR-1998; 98US-0080194P.
PR 01-APR-1998; 98US-0080327P.

PR 01-APR-1998; 98US-0080333P.
PR 08-APR-1998; 98US-0081049P.
PR 08-APR-1998; 98US-0081070P.
PR 09-APR-1998; 98US-0081195P.
PR 15-APR-1998; 98US-0081838P.
PR 21-APR-1998; 98US-0082568P.
PR 21-APR-1998; 98US-0082569P.
PR 22-APR-1998; 98US-0082704P.
PR 22-APR-1998; 98US-0082797P.
PR 28-APR-1998; 98US-0083322P.
PR 29-APR-1998; 98US-0083495P.
PR 29-APR-1998; 98US-0083496P.
PR 29-APR-1998; 98US-0083499P.
PR 29-APR-1998; 98US-0083559P.
PR 05-MAY-1998; 98US-0084366P.
PR 06-MAY-1998; 98US-0084414P.
PR 07-MAY-1998; 98US-0084639P.
PR 07-MAY-1998; 98US-0084640P.
PR 07-MAY-1998; 98US-0084643P.
PR 15-MAY-1998; 98US-0085579P.
PR 15-MAY-1998; 98US-0085580P.
PR 15-MAY-1998; 98US-0085582P.
PR 15-MAY-1998; 98US-0085700P.
PR 18-MAY-1998; 98US-0086023P.
PR 22-MAY-1998; 98US-0086392P.
PR 22-MAY-1998; 98US-0086486P.
PR 28-MAY-1998; 98US-0087098P.
PR 28-MAY-1998; 98US-0087208P.
PR 02-JUN-1998; 98US-0087609P.
PR 02-JUN-1998; 98US-0087759P.
PR 03-JUN-1998; 98US-0087827P.
PR 04-JUN-1998; 98US-0088025P.
PR 04-JUN-1998; 98US-0088028P.
PR 04-JUN-1998; 98US-0088029P.
PR 04-JUN-1998; 98US-0088033P.
PR 04-JUN-1998; 98US-0088326P.
PR 05-JUN-1998; 98US-0088167P.
PR 05-JUN-1998; 98US-0088202P.
PR 05-JUN-1998; 98US-0088212P.
PR 05-JUN-1998; 98US-0088217P.
PR 09-JUN-1998; 98US-0088655P.
PR 10-JUN-1998; 98US-0088722P.
PR 10-JUN-1998; 98US-0088738P.
PR 10-JUN-1998; 98US-0088740P.
PR 10-JUN-1998; 98US-0088811P.
PR 10-JUN-1998; 98US-0088824P.
PR 10-JUN-1998; 98US-0088825P.
PR 10-JUN-1998; 98US-0088826P.
PR 11-JUN-1998; 98US-0088861P.
PR 11-JUN-1998; 98US-0088863P.
PR 11-JUN-1998; 98US-0088876P.
PR 12-JUN-1998; 98US-0089090P.
PR 12-JUN-1998; 98US-0089105P.
PR 16-JUN-1998; 98US-0089512P.
PR 16-JUN-1998; 98US-0089514P.
PR 17-JUN-1998; 98US-0089538P.
PR 17-JUN-1998; 98US-0089598P.
PR 17-JUN-1998; 98US-0089653P.
PR 18-JUN-1998; 98US-0089908P.
PR 19-JUN-1998; 98US-0089952P.
PR 22-JUN-1998; 98US-0090246P.
PR 22-JUN-1998; 98US-0090252P.
PR 22-JUN-1998; 98US-0090254P.
PR 24-JUN-1998; 98US-0090429P.
PR 24-JUN-1998; 98US-0090435P.
PR 24-JUN-1998; 98US-0090444P.
PR 24-JUN-1998; 98US-0090461P.
PR 24-JUN-1998; 98US-0090535P.
PR 24-JUN-1998; 98US-0090540P.
PR 25-JUN-1998; 98US-0090676P.
PR 25-JUN-1998; 98US-0090678P.
PR 25-JUN-1998; 98US-0090688P.
PR 25-JUN-1998; 98US-0090690P.
PR 25-JUN-1998; 98US-0090694P.
PR 25-JUN-1998; 98US-0090695P.
PR 25-JUN-1998; 98US-0090696P.
PR 26-JUN-1998; 98US-00105413.
PR 26-JUN-1998; 98US-0090862P.
PR 26-JUN-1998; 98US-0090863P.
PR 26-JUN-1998; 98US-0091010P.
PR 01-JUL-1998; 98US-0091359P.
PR 01-JUL-1998; 98US-0091544P.
PR 02-JUL-1998; 98US-0091478P.
PR 02-JUL-1998; 98US-0091486P.
PR 02-JUL-1998; 98US-0091626P.
PR 02-JUL-1998; 98US-0091628P.
PR 02-JUL-1998; 98US-0091632P.
PR 24-JUL-1998; 98US-0094006P.
PR 04-AUG-1998; 98US-0095282P.
PR 10-AUG-1998; 98US-0095998P.
PR 10-AUG-1998; 98US-0096012P.
PR 17-AUG-1998; 98US-0096757P.
PR 17-AUG-1998; 98US-0096766P.
PR 17-AUG-1998; 98US-0096867P.
PR 17-AUG-1998; 98US-0096891P.
PR 17-AUG-1998; 98US-0096897P.
PR 18-AUG-1998; 98US-0096949P.
PR 18-AUG-1998; 98US-0096959P.
PR 18-AUG-1998; 98US-0097022P.
PR 26-AUG-1998; 98US-0097952P.
PR 26-AUG-1998; 98US-0097954P.
PR 26-AUG-1998; 98US-0097955P.
PR 26-AUG-1998; 98US-0097971P.
PR 26-AUG-1998; 98US-0097974P.
PR 26-AUG-1998; 98US-0098014P.
PR 01-SEP-1998; 98US-0098716P.
PR 01-SEP-1998; 98US-0098723P.
PR 02-SEP-1998; 98US-0098803P.
PR 02-SEP-1998; 98US-0098821P.
PR 02-SEP-1998; 98US-0098843P.
PR 09-SEP-1998; 98US-0099602P.
PR 10-SEP-1998; 98US-0099741P.
PR 10-SEP-1998; 98US-0099754P.
PR 10-SEP-1998; 98US-0099763P.
PR 10-SEP-1998; 98US-0099812P.
PR 15-SEP-1998; 98US-0100388P.
PR 16-SEP-1998; 98US-0100662P.
PR 16-SEP-1998; 98US-0100664P.
PR 16-SEP-1998; 98US-0101751P.
PR 16-SEP-1998; 98WO-US019330.
PR 17-SEP-1998; 98US-0100683P.
PR 17-SEP-1998; 98US-0100684P.
PR 17-SEP-1998; 98US-0100919P.
PR 17-SEP-1998; 98US-0100930P.
PR 18-SEP-1998; 98US-0100849P.
PR 18-SEP-1998; 98US-0101014P.
PR 18-SEP-1998; 98US-0101068P.
PR 23-SEP-1998; 98US-0101471P.
PR 23-SEP-1998; 98US-0101472P.
PR 23-SEP-1998; 98US-0101475P.
PR 23-SEP-1998; 98US-0101477P.
PR 24-SEP-1998; 98US-0101738P.
PR 24-SEP-1998; 98US-0101739P.
PR 24-SEP-1998; 98US-0101743P.
PR 24-SEP-1998; 98US-0101922P.
PR 25-SEP-1998; 98US-0101786P.
PR 29-SEP-1998; 98US-0102207P.
PR 29-SEP-1998; 98US-0102240P.
PR 29-SEP-1998; 98US-0102330P.
PR 29-SEP-1998; 98US-0102331P.
PR 30-SEP-1998; 98US-0102487P.
PR 30-SEP-1998; 98US-0102570P.
PR 30-SEP-1998; 98US-0102571P.
PR 01-OCT-1998; 98US-0102684P.
PR 01-OCT-1998; 98US-0102687P.

```
  Query Match             98.7%;  Score 1101;  DB 6;  Length 382;
  Best Local Similarity   96.3%;  Pred. No. 5.2e-103;
  Matches  206;  Conservative    0;  Mismatches    0;  Indels    8;  Gaps    1;

Qy          1 ATGRLLS--------GQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 52
              |||||||        ||||||||||||||||||||||||||||||||||||||||||||
Db         22 ATGRLLSASDLDLRGGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 81

Qy         53 ESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVD 112
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 ESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVD 141

Qy        113 EPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGE 172
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 EPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGE 201

Qy        173 ETELTTPVLPEETQEEDAKKTFKESREAALNLAY 206
              |||||||||||||||||||||||||||||||||
Db        202 ETELTTPVLPEETQEEDAKKTFKESREAALNLAY 235


RESULT 44
ABO18815
ID   ABO18815 standard; protein; 382 AA.
XX
AC   ABO18815;
XX
DT   30-AUG-2003  (first entry)
XX
DE   Human secreted/transmembrane protein (PRO) #10.
XX
KW   Human; secreted and transmembrane protein; PRO; TNF-alpha;
KW   tumour necrosis factor alpha; chondrocyte cell; tumour; gene therapy;
KW   tissue typing; adrenal tumour; lung tumour; colon tumour; breast tumour;
KW   prostate tumour; rectal tumour; cervical tumour; liver tumour.
XX
OS   Homo sapiens.
XX
PN   US2003044925-A1.
XX
PD   06-MAR-2003.
XX
PF   25-JUN-2002; 2002US-00180560.
XX
PR   18-SEP-1997;   97US-0059263P.
PR   18-SEP-1997;   97US-0059266P.
PR   17-OCT-1997;   97US-0062250P.
PR   21-OCT-1997;   97US-0063486P.
PR   24-OCT-1997;   97US-0063120P.
PR   24-OCT-1997;   97US-0063121P.
PR   28-OCT-1997;   97US-0063540P.
PR   28-OCT-1997;   97US-0063541P.
PR   28-OCT-1997;   97US-0063544P.
PR   28-OCT-1997;   97US-0063564P.
PR   29-OCT-1997;   97US-0063734P.
PR   31-OCT-1997;   97US-0063870P.
PR   31-OCT-1997;   97US-0064103P.
PR   13-NOV-1997;   97US-0065311P.
PR   21-NOV-1997;   97US-0066120P.
PR   24-NOV-1997;   97US-0066466P.
PR   24-NOV-1997;   97US-0066772P.
PR   11-DEC-1997;   97US-0069335P.
PR   12-DEC-1997;   97US-0069425P.
PR   17-DEC-1997;   97US-0069870P.
PR   18-DEC-1997;   97US-0068017P.
PR   10-MAR-1998;   98US-0077450P.
PR   11-MAR-1998;   98US-0077632P.
PR   11-MAR-1998;   98US-0077649P.
PR   20-MAR-1998;   98US-0078886P.
PR   20-MAR-1998;   98US-0078939P.
PR   27-MAR-1998;   98US-0079664P.
PR   27-MAR-1998;   98US-0079786P.
PR   31-MAR-1998;   98US-0080107P.
PR   31-MAR-1998;   98US-0080194P.
PR   01-APR-1998;   98US-0080327P.
PR   01-APR-1998;   98US-0080333P.
PR   08-APR-1998;   98US-0081049P.
PR   08-APR-1998;   98US-0081070P.
PR   09-APR-1998;   98US-0081195P.
PR   15-APR-1998;   98US-0081838P.
PR   21-APR-1998;   98US-0082568P.
PR   21-APR-1998;   98US-0082569P.
PR   22-APR-1998;   98US-0082704P.
PR   22-APR-1998;   98US-0082797P.
PR   28-APR-1998;   98US-0083322P.
PR   29-APR-1998;   98US-0083495P.
PR   29-APR-1998;   98US-0083496P.
PR   29-APR-1998;   98US-0083499P.
PR   29-APR-1998;   98US-0083559P.
PR   05-MAY-1998;   98US-0084366P.
PR   06-MAY-1998;   98US-0084414P.
PR   07-MAY-1998;   98US-0084639P.
PR   07-MAY-1998;   98US-0084640P.
PR   07-MAY-1998;   98US-0084643P.
PR   15-MAY-1998;   98US-0085579P.
PR   15-MAY-1998;   98US-0085580P.
PR   15-MAY-1998;   98US-0085582P.
PR   15-MAY-1998;   98US-0085700P.
PR   18-MAY-1998;   98US-0086023P.
PR   22-MAY-1998;   98US-0086392P.
PR   22-MAY-1998;   98US-0086486P.
PR   28-MAY-1998;   98US-0087098P.
PR   28-MAY-1998;   98US-0087208P.
PR   02-JUN-1998;   98US-0087609P.
PR   02-JUN-1998;   98US-0087759P.
PR   03-JUN-1998;   98US-0087827P.
PR   04-JUN-1998;   98US-0088025P.
PR   04-JUN-1998;   98US-0088028P.
PR   04-JUN-1998;   98US-0088029P.
PR   04-JUN-1998;   98US-0088033P.
PR   04-JUN-1998;   98US-0088326P.
PR   05-JUN-1998;   98US-0088167P.
PR   05-JUN-1998;   98US-0088202P.
PR   05-JUN-1998;   98US-0088212P.
PR   05-JUN-1998;   98US-0088217P.
PR   09-JUN-1998;   98US-0088655P.
PR   10-JUN-1998;   98US-0088722P.
PR   10-JUN-1998;   98US-0088738P.
PR   10-JUN-1998;   98US-0088740P.
PR   10-JUN-1998;   98US-0088811P.
PR   10-JUN-1998;   98US-0088824P.
PR   10-JUN-1998;   98US-0088825P.
PR   10-JUN-1998;   98US-0088826P.
PR   11-JUN-1998;   98US-0088861P.
PR   11-JUN-1998;   98US-0088863P.
PR   11-JUN-1998;   98US-0088876P.
PR   12-JUN-1998;   98US-0089090P.
PR   12-JUN-1998;   98US-0089105P.
PR   16-JUN-1998;   98US-0089512P.
PR   16-JUN-1998;   98US-0089514P.
PR   17-JUN-1998;   98US-0089538P.
PR   17-JUN-1998;   98US-0089598P.
PR   17-JUN-1998;   98US-0089653P.
PR   18-JUN-1998;   98US-0089908P.
PR   19-JUN-1998;   98US-0089952P.
PR   22-JUN-1998;   98US-0090246P.
PR   22-JUN-1998;   98US-0090252P.
PR   22-JUN-1998;   98US-0090254P.
PR   24-JUN-1998;   98US-0090429P.
PR   24-JUN-1998;   98US-0090435P.
PR   24-JUN-1998;   98US-0090444P.
PR   24-JUN-1998;   98US-0090461P.
PR   24-JUN-1998;   98US-0090535P.
PR   24-JUN-1998;   98US-0090540P.
PR   25-JUN-1998;   98US-0090676P.
PR   25-JUN-1998;   98US-0090678P.
```

PR 25-JUN-1998; 98US-0090688P.
PR 25-JUN-1998; 98US-0090690P.
PR 25-JUN-1998; 98US-0090694P.
PR 25-JUN-1998; 98US-0090695P.
PR 25-JUN-1998; 98US-0090696P.
PR 26-JUN-1998; 98US-00105413.
PR 26-JUN-1998; 98US-0090862P.
PR 26-JUN-1998; 98US-0090863P.
PR 26-JUN-1998; 98US-0091010P.
PR 01-JUL-1998; 98US-0091359P.
PR 01-JUL-1998; 98US-0091544P.
PR 02-JUL-1998; 98US-0091478P.
PR 02-JUL-1998; 98US-0091486P.
PR 02-JUL-1998; 98US-0091626P.
PR 02-JUL-1998; 98US-0091628P.
PR 02-JUL-1998; 98US-0091632P.
PR 24-JUL-1998; 98US-0094006P.
PR 04-AUG-1998; 98US-0095282P.
PR 10-AUG-1998; 98US-0095998P.
PR 10-AUG-1998; 98US-0096012P.
PR 17-AUG-1998; 98US-0096757P.
PR 17-AUG-1998; 98US-0096766P.
PR 17-AUG-1998; 98US-0096867P.
PR 17-AUG-1998; 98US-0096891P.
PR 17-AUG-1998; 98US-0096897P.
PR 18-AUG-1998; 98US-0096949P.
PR 18-AUG-1998; 98US-0096959P.
PR 18-AUG-1998; 98US-0097022P.
PR 26-AUG-1998; 98US-0097952P.
PR 26-AUG-1998; 98US-0097954P.
PR 26-AUG-1998; 98US-0097955P.
PR 26-AUG-1998; 98US-0097971P.
PR 26-AUG-1998; 98US-0097974P.
PR 26-AUG-1998; 98US-0098014P.
PR 01-SEP-1998; 98US-0098716P.
PR 01-SEP-1998; 98US-0098723P.
PR 02-SEP-1998; 98US-0098803P.
PR 02-SEP-1998; 98US-0098821P.
PR 02-SEP-1998; 98US-0098843P.
PR 09-SEP-1998; 98US-0099602P.
PR 10-SEP-1998; 98US-0099741P.
PR 10-SEP-1998; 98US-0099754P.
PR 10-SEP-1998; 98US-0099763P.
PR 10-SEP-1998; 98US-0099812P.
PR 15-SEP-1998; 98US-0100388P.
PR 16-SEP-1998; 98US-0100662P.
PR 16-SEP-1998; 98US-0100664P.
PR 16-SEP-1998; 98US-0101751P.
PR 16-SEP-1998; 98WO-US019330.
PR 17-SEP-1998; 98US-0100683P.
PR 17-SEP-1998; 98US-0100684P.
PR 17-SEP-1998; 98US-0100919P.
PR 17-SEP-1998; 98US-0100930P.
PR 18-SEP-1998; 98US-0100849P.
PR 18-SEP-1998; 98US-0101014P.
PR 18-SEP-1998; 98US-0101068P.
PR 23-SEP-1998; 98US-0101471P.
PR 23-SEP-1998; 98US-0101472P.
PR 23-SEP-1998; 98US-0101475P.
PR 23-SEP-1998; 98US-0101477P.
PR 24-SEP-1998; 98US-0101738P.
PR 24-SEP-1998; 98US-0101739P.
PR 24-SEP-1998; 98US-0101743P.
PR 24-SEP-1998; 98US-0101922P.
PR 25-SEP-1998; 98US-0101786P.
PR 29-SEP-1998; 98US-0102207P.
PR 29-SEP-1998; 98US-0102240P.
PR 29-SEP-1998; 98US-0102330P.
PR 29-SEP-1998; 98US-0102331P.
PR 30-SEP-1998; 98US-0102487P.
PR 30-SEP-1998; 98US-0102570P.
PR 30-SEP-1998; 98US-0102571P.
PR 01-OCT-1998; 98US-0102684P.
PR 01-OCT-1998; 98US-0102687P.
PR 02-OCT-1998; 98US-0102965P.
PR 06-OCT-1998; 98US-0103258P.
PR 06-OCT-1998; 98US-0103449P.

Query Match 98.7%; Score 1101; DB 6; Length 382;
Best Local Similarity 96.3%; Pred. No. 5.2e-103;
Matches 206; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

```
Qy          1 ATGRLLS--------GQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 52
              |||||||        ||||||||||||||||||||||||||||||||||||||||||||
Db         22 ATGRLLSASDLDLRGGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 81

Qy         53 ESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVD 112
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 ESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVD 141

Qy        113 EPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGE 172
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 EPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGE 201

Qy        173 ETELTTPVLPEETQEEDAKKTFKESREAALNLAY 206
              |||||||||||||||||||||||||||||||||
Db        202 ETELTTPVLPEETQEEDAKKTFKESREAALNLAY 235
```

RESULT 45
ABR78236
ID ABR78236 standard; protein; 382 AA.
XX
AC ABR78236;
XX
DT 19-SEP-2003 (first entry)
XX
DE Human secreted polypeptide PRO234, SEQ ID NO:20.
XX
KW Human; PRO; secreted protein; transmembrane protein;
KW extracellular domain; tumour necrosis factor-alpha; TNF-alpha;
KW chondrocyte; proliferation; differentiation; cartilage disorder;
KW bone disorder; arthritis; sports injury; cancer; tumour; diagnosis;
KW adrenal tumour; lung; colon; breast; prostate; kidney; rectum; cervix;
KW liver; drug screening; transgenic animal; genetic analysis;
KW antiarthritic; vulnerary; gene therapy.
XX
OS Homo sapiens.
XX
PN US2003054474-A1.
XX
PD 20-MAR-2003.
XX
PF 22-JUL-2002; 2002US-00201530.
XX
PR 22-JUN-1998; 98US-0090254P.
PR 02-JUN-1999; 99WO-US012252.
PR 25-AUG-1999; 99US-00380137.
PR 28-FEB-2001; 2001WO-US006520.
PR 15-JAN-2002; 2002US-00052586.
XX
PA (GETH) GENENTECH INC.
XX
PI Baker KP, Chen J, Desnoyers L, Goddard A, Godowski PJ, Gurney AL;
PI Pan J, Smith V, Watanabe CK, Wood WI, Zhang Z;
XX
DR WPI; 2003-503631/47.
DR N-PSDB; ACF00109.
XX
PT New secreted and transmembrane PRO polypeptides and nucleic acids, useful
PT in gene therapy, or for preparing a medicament for treating a condition
PT that is responsive to the PRO polypeptide or anti-PRO antibody.
XX
PS Claim 11; Fig 20; 700pp; English.
XX
CC The invention relates to human PRO secreted/transmembrane polypeptides

CC (ABR78227-ABR78531) and nucleic acids encoding them (ACF00100-00404). The
CC invention also relates to sequences at least 80% identical to the PRO
CC nucleic acid and polypeptide sequences of the invention, recombinant
CC vectors and host cells comprising a PRO nucleic acid, a method for the
CC recombinant production of a PRO polypeptide, antibodies against a PRO
CC polypeptide, and fusion proteins comprising a PRO polypeptide. Nucleic
CC acids encoding PRO polypeptides of the invention were initially
CC identified via homology screening using consensus sequences based on the
CC extracellular domain sequences from known secreted proteins. Human cDNA
CC libraries containing sequences of interest were identified using
CC oligonucleotides based on the consensus sequences, and cDNA clones were
CC isolated and characterised. The PRO polypeptides are useful for
CC stimulating release of tumour necrosis factor-alpha (TNF-alpha) from
CC human blood and may thus be used in the treatment of conditions in which
CC enhanced TNF-alpha release would be beneficial. They are also useful for
CC stimulating the proliferation or differentiation of chondrocytes and as
CC such may be used in the treatment of various bone and/or cartilage
CC disorders such as arthritis and sports injuries. The PRO polypeptides may
CC be used in a method for detecting the presence of a tumour (e.g., an
CC adrenal tumour, lung tumour, colon tumour, breast tumour, prostate
CC tumour, rectal tumour, cervical tumour or liver tumour) in a mammal. This
CC method involves comparing the level of expression of the PRO polypeptide
CC in test and control samples, where a higher level of expression of PRO
CC polypeptide in the test sample as compared to the control sample is
CC indicative of the presence of a tumour. The PRO polypeptides are
CC additionally useful for in drug screening to identify agonists and
CC antagonists of PRO polypeptides. PRO nucleic acids are useful as
CC hybridisation probes (for isolation of cDNA molecules), in chromosome and
CC gene mapping, in the generation of antisense RNA and DNA and in gene
CC therapy. The nucleic acids can also be used for mapping genes encoding
CC PRO polypeptides, for genetic analysis of individuals with genetic
CC disorders, and for generating either transgenic animals or knock-out
CC animals which are useful in the development and screening of
CC therapeutically useful compounds. Sequences ABR78227-ABR78531 represent
CC the human PRO secreted/transmembrane polypeptides of the invention. Note:
CC The sequence data for this patent is also available in electronic format
CC from USPTO at seqdata.uspto.gov/sequence.html
XX
SQ Sequence 382 AA;

```
  Query Match             98.7%;  Score 1101;  DB 6;  Length 382;
  Best Local Similarity   96.3%;  Pred. No. 5.2e-103;
  Matches  206;  Conservative    0;  Mismatches    0;  Indels    8;  Gaps    1;

Qy          1 ATGRLLS--------GQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 52
              |||||||        ||||||||||||||||||||||||||||||||||||||||||||
Db         22 ATGRLLSASDLDLRGGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 81

Qy         53 ESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVD 112
              |||||||||.||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 ESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVD 141

Qy        113 EPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGE 172
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 EPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGE 201

Qy        173 ETELTTPVLPEETQEEDAKKTFKESREAALNLAY 206
              |||||||||||||||||||||||||||||||||
Db        202 ETELTTPVLPEETQEEDAKKTFKESREAALNLAY 235
```

RESULT 46
ABU84972
ID ABU84972 standard; protein; 382 AA.
XX
AC ABU84972;
XX
DT 30-JUN-2003 (first entry)
XX
DE Novel human secreted and transmembrane protein PRO234.
XX
KW Human; secreted and transmembrane protein; PRO; cytostatic; gene therapy;
KW chondrocyte stimulator; chromosome mapping; gene mapping;
KW transgenic animal; knock-out animal; tumour.
XX
OS Homo sapiens.
XX
PN US2003032114-A1.
XX
PD 13-FEB-2003.
XX
PF 20-JUN-2002; 2002US-00176919.
XX
PR 18-SEP-1997; 97US-0059263P.
PR 18-SEP-1997; 97US-0059266P.
PR 17-OCT-1997; 97US-0062250P.
PR 21-OCT-1997; 97US-0063486P.
PR 24-OCT-1997; 97US-0063120P.
PR 24-OCT-1997; 97US-0063121P.
PR 28-OCT-1997; 97US-0063540P.
PR 28-OCT-1997; 97US-0063541P.
PR 28-OCT-1997; 97US-0063544P.
PR 28-OCT-1997; 97US-0063564P.
PR 29-OCT-1997; 97US-0063734P.
PR 31-OCT-1997; 97US-0063870P.
PR 31-OCT-1997; 97US-0064103P.
PR 13-NOV-1997; 97US-0065311P.
PR 21-NOV-1997; 97US-0066120P.
PR 24-NOV-1997; 97US-0066466P.
PR 24-NOV-1997; 97US-0066772P.
PR 11-DEC-1997; 97US-0069335P.
PR 12-DEC-1997; 97US-0069425P.
PR 17-DEC-1997; 97US-0069870P.
PR 18-DEC-1997; 97US-0068017P.
PR 10-MAR-1998; 98US-0077450P.
PR 11-MAR-1998; 98US-0077632P.
PR 11-MAR-1998; 98US-0077649P.
PR 20-MAR-1998; 98US-0078886P.
PR 20-MAR-1998; 98US-0078939P.
PR 27-MAR-1998; 98US-0079664P.
PR 27-MAR-1998; 98US-0079786P.
PR 31-MAR-1998; 98US-0080107P.
PR 31-MAR-1998; 98US-0080194P.
PR 01-APR-1998; 98US-0080327P.
PR 01-APR-1998; 98US-0080333P.
PR 08-APR-1998; 98US-0081049P.
PR 08-APR-1998; 98US-0081070P.
PR 09-APR-1998; 98US-0081195P.
PR 15-APR-1998; 98US-0081838P.
PR 21-APR-1998; 98US-0082568P.
PR 21-APR-1998; 98US-0082569P.
PR 22-APR-1998; 98US-0082704P.
PR 22-APR-1998; 98US-0082797P.
PR 28-APR-1998; 98US-0083322P.
PR 29-APR-1998; 98US-0083495P.
PR 29-APR-1998; 98US-0083496P.
PR 29-APR-1998; 98US-0083499P.
PR 29-APR-1998; 98US-0083559P.
PR 05-MAY-1998; 98US-0084366P.
PR 06-MAY-1998; 98US-0084414P.
PR 07-MAY-1998; 98US-0084639P.
PR 07-MAY-1998; 98US-0084640P.
PR 07-MAY-1998; 98US-0084643P.
PR 15-MAY-1998; 98US-0085579P.
PR 15-MAY-1998; 98US-0085580P.
PR 15-MAY-1998; 98US-0085582P.
PR 15-MAY-1998; 98US-0085700P.
PR 18-MAY-1998; 98US-0086023P.
PR 22-MAY-1998; 98US-0086392P.
PR 22-MAY-1998; 98US-0086486P.
PR 28-MAY-1998; 98US-0087098P.
PR 28-MAY-1998; 98US-0087208P.
PR 02-JUN-1998; 98US-0087609P.
PR 02-JUN-1998; 98US-0087759P.
PR 03-JUN-1998; 98US-0087827P.
PR 04-JUN-1998; 98US-0088025P.

PR 04-JUN-1998; 98US-0088028P.
PR 04-JUN-1998; 98US-0088029P.
PR 04-JUN-1998; 98US-0088033P.
PR 04-JUN-1998; 98US-0088326P.
PR 05-JUN-1998; 98US-0088167P.
PR 05-JUN-1998; 98US-0088202P.
PR 05-JUN-1998; 98US-0088212P.
PR 05-JUN-1998; 98US-0088217P.
PR 09-JUN-1998; 98US-0088655P.
PR 10-JUN-1998; 98US-0088722P.
PR 10-JUN-1998; 98US-0088738P.
PR 10-JUN-1998; 98US-0088740P.
PR 10-JUN-1998; 98US-0088811P.
PR 10-JUN-1998; 98US-0088824P.
PR 10-JUN-1998; 98US-0088825P.
PR 10-JUN-1998; 98US-0088826P.
PR 11-JUN-1998; 98US-0088861P.
PR 11-JUN-1998; 98US-0088863P.
PR 11-JUN-1998; 98US-0088876P.
PR 12-JUN-1998; 98US-0089090P.
PR 12-JUN-1998; 98US-0089105P.
PR 16-JUN-1998; 98US-0089512P.
PR 16-JUN-1998; 98US-0089514P.
PR 17-JUN-1998; 98US-0089538P.
PR 17-JUN-1998; 98US-0089598P.
PR 17-JUN-1998; 98US-0089653P.
PR 18-JUN-1998; 98US-0089908P.
PR 19-JUN-1998; 98US-0089952P.
PR 22-JUN-1998; 98US-0090246P.
PR 22-JUN-1998; 98US-0090252P.
PR 22-JUN-1998; 98US-0090254P.
PR 24-JUN-1998; 98US-0090429P.
PR 24-JUN-1998; 98US-0090435P.
PR 24-JUN-1998; 98US-0090444P.
PR 24-JUN-1998; 98US-0090461P.
PR 24-JUN-1998; 98US-0090535P.
PR 24-JUN-1998; 98US-0090540P.
PR 25-JUN-1998; 98US-0090676P.
PR 25-JUN-1998; 98US-0090678P.
PR 25-JUN-1998; 98US-0090688P.
PR 25-JUN-1998; 98US-0090690P.
PR 25-JUN-1998; 98US-0090694P.
PR 25-JUN-1998; 98US-0090695P.
PR 25-JUN-1998; 98US-0090696P.
PR 26-JUN-1998; 98US-00105413.
PR 26-JUN-1998; 98US-0090862P.
PR 26-JUN-1998; 98US-0090863P.
PR 26-JUN-1998; 98US-0091010P.
PR 01-JUL-1998; 98US-0091359P.
PR 01-JUL-1998; 98US-0091544P.
PR 02-JUL-1998; 98US-0091478P.
PR 02-JUL-1998; 98US-0091486P.
PR 02-JUL-1998; 98US-0091626P.
PR 02-JUL-1998; 98US-0091628P.
PR 02-JUL-1998; 98US-0091632P.
PR 24-JUL-1998; 98US-0094006P.
PR 04-AUG-1998; 98US-0095282P.
PR 10-AUG-1998; 98US-0095998P.
PR 10-AUG-1998; 98US-0096012P.
PR 17-AUG-1998; 98US-0096757P.
PR 17-AUG-1998; 98US-0096766P.
PR 17-AUG-1998; 98US-0096867P.
PR 17-AUG-1998; 98US-0096891P.
PR 17-AUG-1998; 98US-0096897P.
PR 18-AUG-1998; 98US-0096949P.
PR 18-AUG-1998; 98US-0096959P.
PR 18-AUG-1998; 98US-0097022P.
PR 26-AUG-1998; 98US-0097952P.
PR 26-AUG-1998; 98US-0097954P.
PR 26-AUG-1998; 98US-0097955P.
PR 26-AUG-1998; 98US-0097971P.
PR 26-AUG-1998; 98US-0097974P.
PR 26-AUG-1998; 98US-0098014P.
PR 01-SEP-1998; 98US-0098716P.
PR 01-SEP-1998; 98US-0098723P.
PR 02-SEP-1998; 98US-0098803P.
PR 02-SEP-1998; 98US-0098821P.
PR 02-SEP-1998; 98US-0098843P.
PR 09-SEP-1998; 98US-0099602P.
PR 10-SEP-1998; 98US-0099741P.
PR 10-SEP-1998; 98US-0099754P.
PR 10-SEP-1998; 98US-0099763P.
PR 10-SEP-1998; 98US-0099812P.
PR 15-SEP-1998; 98US-0100388P.
PR 16-SEP-1998; 98US-0100662P.
PR 16-SEP-1998; 98US-0100664P.
PR 16-SEP-1998; 98US-0101751P.
PR 16-SEP-1998; 98WO-US019330.
PR 17-SEP-1998; 98US-0100683P.
PR 17-SEP-1998; 98US-0100684P.
PR 17-SEP-1998; 98US-0100919P.
PR 17-SEP-1998; 98US-0100930P.
PR 18-SEP-1998; 98US-0100849P.
PR 18-SEP-1998; 98US-0101014P.
PR 18-SEP-1998; 98US-0101068P.
PR 23-SEP-1998; 98US-0101471P.
PR 23-SEP-1998; 98US-0101472P.
PR 23-SEP-1998; 98US-0101475P.
PR 23-SEP-1998; 98US-0101477P.
PR 24-SEP-1998; 98US-0101738P.
PR 24-SEP-1998; 98US-0101739P.
PR 24-SEP-1998; 98US-0101743P.
PR 24-SEP-1998; 98US-0101922P.
PR 25-SEP-1998; 98US-0101786P.
PR 29-SEP-1998; 98US-0102207P.
PR 29-SEP-1998; 98US-0102240P.
PR 29-SEP-1998; 98US-0102330P.
PR 29-SEP-1998; 98US-0102331P.
PR 30-SEP-1998; 98US-0102487P.
PR 30-SEP-1998; 98US-0102570P.
PR 30-SEP-1998; 98US-0102571P.
PR 01-OCT-1998; 98US-0102684P.
PR 01-OCT-1998; 98US-0102687P.
PR 02-OCT-1998; 98US-0102965P.
PR 06-OCT-1998; 98US-0103258P.
PR 06-OCT-1998; 98US-0103449P.
PR 07-OCT-1998; 98US-00168978.

Query Match 98.7%; Score 1101; DB 6; Length 382;
Best Local Similarity 96.3%; Pred. No. 5.2e-103;
Matches 206; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

```
Qy          1 ATGRLLS--------GQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 52
              |||||||        ||||||||||||||||||||||||||||||||||||||||||||
Db         22 ATGRLLSASDLDLRGGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 81

Qy         53 ESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVD 112
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 ESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVD 141

Qy        113 EPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGE 172
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 EPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGE 201

Qy        173 ETELTTPVLPEETQEEDAKKTFKESREAALNLAY 206
              |||||||||||||||||||||||||||||||||
Db        202 ETELTTPVLPEETQEEDAKKTFKESREAALNLAY 235
```

RESULT 47
ABO00111
ID ABO00111 standard; protein; 382 AA.
XX
AC ABO00111;
XX
DT 06-AUG-2003 (first entry)

XX
DE Novel human secreted and transmembrane protein PRO234.
XX
KW Human; gene therapy; tumour necrosis factor alpha; TNF-alpha;
KW chondrocyte stimulation; tumour; tissue typing.
XX
OS Homo sapiens.
XX
PN US2003032101-A1.
XX
PD 13-FEB-2003.
XX
PF 17-JUN-2002; 2002US-00173695.
XX
PR 18-SEP-1997; 97US-0059263P.
PR 18-SEP-1997; 97US-0059266P.
PR 17-OCT-1997; 97US-0062250P.
PR 21-OCT-1997; 97US-0063486P.
PR 24-OCT-1997; 97US-0063120P.
PR 24-OCT-1997; 97US-0063121P.
PR 28-OCT-1997; 97US-0063540P.
PR 28-OCT-1997; 97US-0063541P.
PR 28-OCT-1997; 97US-0063544P.
PR 28-OCT-1997; 97US-0063564P.
PR 29-OCT-1997; 97US-0063734P.
PR 31-OCT-1997; 97US-0063870P.
PR 31-OCT-1997; 97US-0064103P.
PR 13-NOV-1997; 97US-0065311P.
PR 21-NOV-1997; 97US-0066120P.
PR 24-NOV-1997; 97US-0066466P.
PR 24-NOV-1997; 97US-0066772P.
PR 11-DEC-1997; 97US-0069335P.
PR 12-DEC-1997; 97US-0069425P.
PR 17-DEC-1997; 97US-0069870P.
PR 18-DEC-1997; 97US-0068017P.
PR 10-MAR-1998; 98US-0077450P.
PR 11-MAR-1998; 98US-0077632P.
PR 11-MAR-1998; 98US-0077649P.
PR 20-MAR-1998; 98US-0078886P.
PR 20-MAR-1998; 98US-0078939P.
PR 27-MAR-1998; 98US-0079664P.
PR 27-MAR-1998; 98US-0079786P.
PR 31-MAR-1998; 98US-0080107P.
PR 31-MAR-1998; 98US-0080194P.
PR 01-APR-1998; 98US-0080327P.
PR 01-APR-1998; 98US-0080333P.
PR 08-APR-1998; 98US-0081049P.
PR 08-APR-1998; 98US-0081070P.
PR 09-APR-1998; 98US-0081195P.
PR 15-APR-1998; 98US-0081838P.
PR 21-APR-1998; 98US-0082568P.
PR 21-APR-1998; 98US-0082569P.
PR 22-APR-1998; 98US-0082704P.
PR 22-APR-1998; 98US-0082797P.
PR 28-APR-1998; 98US-0083322P.
PR 29-APR-1998; 98US-0083495P.
PR 29-APR-1998; 98US-0083496P.
PR 29-APR-1998; 98US-0083499P.
PR 29-APR-1998; 98US-0083559P.
PR 05-MAY-1998; 98US-0084366P.
PR 06-MAY-1998; 98US-0084414P.
PR 07-MAY-1998; 98US-0084639P.
PR 07-MAY-1998; 98US-0084640P.
PR 07-MAY-1998; 98US-0084643P.
PR 15-MAY-1998; 98US-0085579P.
PR 15-MAY-1998; 98US-0085580P.
PR 15-MAY-1998; 98US-0085582P.
PR 15-MAY-1998; 98US-0085700P.
PR 18-MAY-1998; 98US-0086023P.
PR 22-MAY-1998; 98US-0086392P.
PR 22-MAY-1998; 98US-0086486P.
PR 28-MAY-1998; 98US-0087098P.
PR 28-MAY-1998; 98US-0087208P.
PR 02-JUN-1998; 98US-0087609P.
PR 02-JUN-1998; 98US-0087759P.
PR 03-JUN-1998; 98US-0087827P.
PR 04-JUN-1998; 98US-0088025P.
PR 04-JUN-1998; 98US-0088028P.
PR 04-JUN-1998; 98US-0088029P.
PR 04-JUN-1998; 98US-0088033P.
PR 04-JUN-1998; 98US-0088326P.
PR 05-JUN-1998; 98US-0088167P.
PR 05-JUN-1998; 98US-0088202P.
PR 05-JUN-1998; 98US-0088212P.
PR 05-JUN-1998; 98US-0088217P.
PR 09-JUN-1998; 98US-0088655P.
PR 10-JUN-1998; 98US-0088722P.
PR 10-JUN-1998; 98US-0088738P.
PR 10-JUN-1998; 98US-0088740P.
PR 10-JUN-1998; 98US-0088811P.
PR 10-JUN-1998; 98US-0088824P.
PR 10-JUN-1998; 98US-0088825P.
PR 10-JUN-1998; 98US-0088826P.
PR 11-JUN-1998; 98US-0088861P.
PR 11-JUN-1998; 98US-0088863P.
PR 11-JUN-1998; 98US-0088876P.
PR 12-JUN-1998; 98US-0089090P.
PR 12-JUN-1998; 98US-0089105P.
PR 16-JUN-1998; 98US-0089512P.
PR 16-JUN-1998; 98US-0089514P.
PR 17-JUN-1998; 98US-0089538P.
PR 17-JUN-1998; 98US-0089598P.
PR 17-JUN-1998; 98US-0089653P.
PR 18-JUN-1998; 98US-0089908P.
PR 19-JUN-1998; 98US-0089952P.
PR 22-JUN-1998; 98US-0090246P.
PR 22-JUN-1998; 98US-0090252P.
PR 22-JUN-1998; 98US-0090254P.
PR 24-JUN-1998; 98US-0090429P.
PR 24-JUN-1998; 98US-0090435P.
PR 24-JUN-1998; 98US-0090444P.
PR 24-JUN-1998; 98US-0090461P.
PR 24-JUN-1998; 98US-0090535P.
PR 24-JUN-1998; 98US-0090540P.
PR 25-JUN-1998; 98US-0090676P.
PR 25-JUN-1998; 98US-0090678P.
PR 25-JUN-1998; 98US-0090688P.
PR 25-JUN-1998; 98US-0090690P.
PR 25-JUN-1998; 98US-0090694P.
PR 25-JUN-1998; 98US-0090695P.
PR 25-JUN-1998; 98US-0090696P.
PR 26-JUN-1998; 98US-00105413.
PR 26-JUN-1998; 98US-0090862P.
PR 26-JUN-1998; 98US-0090863P.
PR 26-JUN-1998; 98US-0091010P.
PR 01-JUL-1998; 98US-0091359P.
PR 01-JUL-1998; 98US-0091544P.
PR 02-JUL-1998; 98US-0091478P.
PR 02-JUL-1998; 98US-0091486P.
PR 02-JUL-1998; 98US-0091626P.
PR 02-JUL-1998; 98US-0091628P.
PR 02-JUL-1998; 98US-0091632P.
PR 24-JUL-1998; 98US-0094006P.
PR 04-AUG-1998; 98US-0095282P.
PR 10-AUG-1998; 98US-0095998P.
PR 10-AUG-1998; 98US-0096012P.
PR 17-AUG-1998; 98US-0096757P.
PR 17-AUG-1998; 98US-0096766P.
PR 17-AUG-1998; 98US-0096867P.
PR 17-AUG-1998; 98US-0096891P.
PR 17-AUG-1998; 98US-0096897P.
PR 18-AUG-1998; 98US-0096949P.
PR 18-AUG-1998; 98US-0096959P.
PR 18-AUG-1998; 98US-0097022P.
PR 26-AUG-1998; 98US-0097952P.
PR 26-AUG-1998; 98US-0097954P.

PR 26-AUG-1998; 98US-0097955P.
PR 26-AUG-1998; 98US-0097971P.
PR 26-AUG-1998; 98US-0097974P.
PR 26-AUG-1998; 98US-0098014P.
PR 01-SEP-1998; 98US-0098716P.
PR 01-SEP-1998; 98US-0098723P.
PR 02-SEP-1998; 98US-0098803P.
PR 02-SEP-1998; 98US-0098821P.
PR 02-SEP-1998; 98US-0098843P.
PR 09-SEP-1998; 98US-0099602P.
PR 10-SEP-1998; 98US-0099741P.
PR 10-SEP-1998; 98US-0099754P.
PR 10-SEP-1998; 98US-0099763P.
PR 10-SEP-1998; 98US-0099812P.
PR 15-SEP-1998; 98US-0100388P.
PR 16-SEP-1998; 98US-0100662P.
PR 16-SEP-1998; 98US-0100664P.
PR 16-SEP-1998; 98US-0101751P.
PR 16-SEP-1998; 98WO-US019330.
PR 17-SEP-1998; 98US-0100683P.
PR 17-SEP-1998; 98US-0100684P.
PR 17-SEP-1998; 98US-0100919P.
PR 17-SEP-1998; 98US-0100930P.
PR 18-SEP-1998; 98US-0100849P.
PR 18-SEP-1998; 98US-0101014P.
PR 18-SEP-1998; 98US-0101068P.
PR 23-SEP-1998; 98US-0101471P.
PR 23-SEP-1998; 98US-0101472P.
PR 23-SEP-1998; 98US-0101475P.
PR 23-SEP-1998; 98US-0101477P.
PR 24-SEP-1998; 98US-0101738P.
PR 24-SEP-1998; 98US-0101739P.
PR 24-SEP-1998; 98US-0101743P.
PR 24-SEP-1998; 98US-0101922P.
PR 25-SEP-1998; 98US-0101786P.
PR 29-SEP-1998; 98US-0102207P.
PR 29-SEP-1998; 98US-0102240P.
PR 29-SEP-1998; 98US-0102330P.
PR 29-SEP-1998; 98US-0102331P.
PR 30-SEP-1998; 98US-0102487P.
PR 30-SEP-1998; 98US-0102570P.
PR 30-SEP-1998; 98US-0102571P.
PR 01-OCT-1998; 98US-0102684P.
PR 01-OCT-1998; 98US-0102687P.
PR 02-OCT-1998; 98US-0102965P.
PR 06-OCT-1998; 98US-0103258P.
PR 06-OCT-1998; 98US-0103449P.
PR 07-OCT-1998; 98US-00168978.
PR 07-OCT-1998; 98US-0103395P.

Query Match 98.7%; Score 1101; DB 6; Length 382;
Best Local Similarity 96.3%; Pred. No. 5.2e-103;
Matches 206; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

```
Qy          1 ATGRLLS--------GQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 52
              |||||||        ||||||||||||||||||||||||||||||||||||||||||||
Db         22 ATGRLLSASDLDLRGGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 81

Qy         53 ESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVD 112
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 ESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVD 141

Qy        113 EPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGE 172
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 EPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGE 201

Qy        173 ETELTTPVLPEETQEEDAKKTFKESREAALNLAY 206
              |||||||||||||||||||||||||||||||||
Db        202 ETELTTPVLPEETQEEDAKKTFKESREAALNLAY 235
```

RESULT 48
ABO11443

ID ABO11443 standard; protein; 382 AA.
XX
AC ABO11443;
XX
DT 26-AUG-2003 (first entry)
XX
DE Human secreted/transmembrane protein (PRO) #10.
XX
KW Human; secreted and transmembrane protein; PRO; TNF-alpha;
KW tumour necrosis factor alpha; chondrocyte cell; tumour; gene therapy;
KW tissue typing; adrenal tumour; lung tumour; colon tumour; breast tumour;
KW prostate tumour; rectal tumour; cervical tumour; liver tumour.
XX
OS Homo sapiens.
XX
PN US2003036124-A1.
XX
PD 20-FEB-2003.
XX
PF 26-JUN-2002; 2002US-00180998.
XX
PR 18-SEP-1997; 97US-0059263P.
PR 18-SEP-1997; 97US-0059266P.
PR 17-OCT-1997; 97US-0062250P.
PR 21-OCT-1997; 97US-0063486P.
PR 24-OCT-1997; 97US-0063120P.
PR 24-OCT-1997; 97US-0063121P.
PR 28-OCT-1997; 97US-0063540P.
PR 28-OCT-1997; 97US-0063541P.
PR 28-OCT-1997; 97US-0063544P.
PR 28-OCT-1997; 97US-0063564P.
PR 29-OCT-1997; 97US-0063734P.
PR 31-OCT-1997; 97US-0063870P.
PR 31-OCT-1997; 97US-0064103P.
PR 13-NOV-1997; 97US-0065311P.
PR 21-NOV-1997; 97US-0066120P.
PR 24-NOV-1997; 97US-0066466P.
PR 24-NOV-1997; 97US-0066772P.
PR 11-DEC-1997; 97US-0069335P.
PR 12-DEC-1997; 97US-0069425P.
PR 17-DEC-1997; 97US-0069870P.
PR 18-DEC-1997; 97US-0068017P.
PR 10-MAR-1998; 98US-0077450P.
PR 11-MAR-1998; 98US-0077632P.
PR 11-MAR-1998; 98US-0077649P.
PR 20-MAR-1998; 98US-0078886P.
PR 20-MAR-1998; 98US-0078939P.
PR 27-MAR-1998; 98US-0079664P.
PR 27-MAR-1998; 98US-0079786P.
PR 31-MAR-1998; 98US-0080107P.
PR 31-MAR-1998; 98US-0080194P.
PR 01-APR-1998; 98US-0080327P.
PR 01-APR-1998; 98US-0080333P.
PR 08-APR-1998; 98US-0081049P.
PR 08-APR-1998; 98US-0081070P.
PR 09-APR-1998; 98US-0081195P.
PR 15-APR-1998; 98US-0081838P.
PR 21-APR-1998; 98US-0082568P.
PR 21-APR-1998; 98US-0082569P.
PR 22-APR-1998; 98US-0082704P.
PR 22-APR-1998; 98US-0082797P.
PR 28-APR-1998; 98US-0083322P.
PR 29-APR-1998; 98US-0083495P.
PR 29-APR-1998; 98US-0083496P.
PR 29-APR-1998; 98US-0083499P.
PR 29-APR-1998; 98US-0083559P.
PR 05-MAY-1998; 98US-0084366P.
PR 06-MAY-1998; 98US-0084414P.
PR 07-MAY-1998; 98US-0084639P.
PR 07-MAY-1998; 98US-0084640P.
PR 07-MAY-1998; 98US-0084643P.
PR 15-MAY-1998; 98US-0085579P.
PR 15-MAY-1998; 98US-0085580P.

PR 15-MAY-1998; 98US-0085582P.
PR 15-MAY-1998; 98US-0085700P.
PR 18-MAY-1998; 98US-0086023P.
PR 22-MAY-1998; 98US-0086392P.
PR 22-MAY-1998; 98US-0086486P.
PR 28-MAY-1998; 98US-0087098P.
PR 28-MAY-1998; 98US-0087208P.
PR 02-JUN-1998; 98US-0087609P.
PR 02-JUN-1998; 98US-0087759P.
PR 03-JUN-1998; 98US-0087827P.
PR 04-JUN-1998; 98US-0088025P.
PR 04-JUN-1998; 98US-0088028P.
PR 04-JUN-1998; 98US-0088029P.
PR 04-JUN-1998; 98US-0088033P.
PR 04-JUN-1998; 98US-0088326P.
PR 05-JUN-1998; 98US-0088167P.
PR 05-JUN-1998; 98US-0088202P.
PR 05-JUN-1998; 98US-0088212P.
PR 05-JUN-1998; 98US-0088217P.
PR 09-JUN-1998; 98US-0088655P.
PR 10-JUN-1998; 98US-0088722P.
PR 10-JUN-1998; 98US-0088738P.
PR 10-JUN-1998; 98US-0088740P.
PR 10-JUN-1998; 98US-0088811P.
PR 10-JUN-1998; 98US-0088824P.
PR 10-JUN-1998; 98US-0088825P.
PR 10-JUN-1998; 98US-0088826P.
PR 11-JUN-1998; 98US-0088861P.
PR 11-JUN-1998; 98US-0088863P.
PR 11-JUN-1998; 98US-0088876P.
PR 12-JUN-1998; 98US-0089090P.
PR 12-JUN-1998; 98US-0089105P.
PR 16-JUN-1998; 98US-0089512P.
PR 16-JUN-1998; 98US-0089514P.
PR 17-JUN-1998; 98US-0089538P.
PR 17-JUN-1998; 98US-0089598P.
PR 17-JUN-1998; 98US-0089653P.
PR 18-JUN-1998; 98US-0089908P.
PR 19-JUN-1998; 98US-0089952P.
PR 22-JUN-1998; 98US-0090246P.
PR 22-JUN-1998; 98US-0090252P.
PR 22-JUN-1998; 98US-0090254P.
PR 24-JUN-1998; 98US-0090429P.
PR 24-JUN-1998; 98US-0090435P.
PR 24-JUN-1998; 98US-0090444P.
PR 24-JUN-1998; 98US-0090461P.
PR 24-JUN-1998; 98US-0090535P.
PR 24-JUN-1998; 98US-0090540P.
PR 25-JUN-1998; 98US-0090676P.
PR 25-JUN-1998; 98US-0090678P.
PR 25-JUN-1998; 98US-0090688P.
PR 25-JUN-1998; 98US-0090690P.
PR 25-JUN-1998; 98US-0090694P.
PR 25-JUN-1998; 98US-0090695P.
PR 25-JUN-1998; 98US-0090696P.
PR 26-JUN-1998; 98US-00105413.
PR 26-JUN-1998; 98US-0090862P.
PR 26-JUN-1998; 98US-0090863P.
PR 26-JUN-1998; 98US-0091010P.
PR 01-JUL-1998; 98US-0091359P.
PR 01-JUL-1998; 98US-0091544P.
PR 02-JUL-1998; 98US-0091478P.
PR 02-JUL-1998; 98US-0091486P.
PR 02-JUL-1998; 98US-0091626P.
PR 02-JUL-1998; 98US-0091628P.
PR 02-JUL-1998; 98US-0091632P.
PR 24-JUL-1998; 98US-0094006P.
PR 04-AUG-1998; 98US-0095282P.
PR 10-AUG-1998; 98US-0095998P.
PR 10-AUG-1998; 98US-0096012P.
PR 17-AUG-1998; 98US-0096757P.
PR 17-AUG-1998; 98US-0096766P.
PR 17-AUG-1998; 98US-0096867P.
PR 17-AUG-1998; 98US-0096891P.
PR 17-AUG-1998; 98US-0096897P.
PR 18-AUG-1998; 98US-0096949P.
PR 18-AUG-1998; 98US-0096959P.
PR 18-AUG-1998; 98US-0097022P.
PR 26-AUG-1998; 98US-0097952P.
PR 26-AUG-1998; 98US-0097954P.
PR 26-AUG-1998; 98US-0097955P.
PR 26-AUG-1998; 98US-0097971P.
PR 26-AUG-1998; 98US-0097974P.
PR 26-AUG-1998; 98US-0098014P.
PR 01-SEP-1998; 98US-0098716P.
PR 01-SEP-1998; 98US-0098723P.
PR 02-SEP-1998; 98US-0098803P.
PR 02-SEP-1998; 98US-0098821P.
PR 02-SEP-1998; 98US-0098843P.
PR 09-SEP-1998; 98US-0099602P.
PR 10-SEP-1998; 98US-0099741P.
PR 10-SEP-1998; 98US-0099754P.
PR 10-SEP-1998; 98US-0099763P.
PR 10-SEP-1998; 98US-0099812P.
PR 15-SEP-1998; 98US-0100388P.
PR 16-SEP-1998; 98US-0100662P.
PR 16-SEP-1998; 98US-0100664P.
PR 16-SEP-1998; 98US-0101751P.
PR 16-SEP-1998; 98WO-US019330.
PR 17-SEP-1998; 98US-0100683P.
PR 17-SEP-1998; 98US-0100684P.
PR 17-SEP-1998; 98US-0100919P.
PR 17-SEP-1998; 98US-0100930P.
PR 18-SEP-1998; 98US-0100849P.
PR 18-SEP-1998; 98US-0101014P.
PR 18-SEP-1998; 98US-0101068P.
PR 23-SEP-1998; 98US-0101471P.
PR 23-SEP-1998; 98US-0101472P.
PR 23-SEP-1998; 98US-0101475P.
PR 23-SEP-1998; 98US-0101477P.
PR 24-SEP-1998; 98US-0101738P.
PR 24-SEP-1998; 98US-0101739P.
PR 24-SEP-1998; 98US-0101743P.
PR 24-SEP-1998; 98US-0101922P.
PR 25-SEP-1998; 98US-0101786P.
PR 29-SEP-1998; 98US-0102207P.
PR 29-SEP-1998; 98US-0102240P.
PR 29-SEP-1998; 98US-0102330P.
PR 29-SEP-1998; 98US-0102331P.
PR 30-SEP-1998; 98US-0102487P.
PR 30-SEP-1998; 98US-0102570P.
PR 30-SEP-1998; 98US-0102571P.
PR 01-OCT-1998; 98US-0102684P.
PR 01-OCT-1998; 98US-0102687P.
PR 02-OCT-1998; 98US-0102965P.
PR 06-OCT-1998; 98US-0103258P.
PR 06-OCT-1998; 98US-0103449P.

Query Match 98.7%; Score 1101; DB 6; Length 382;
Best Local Similarity 96.3%; Pred. No. 5.2e-103;
Matches 206; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

```
Qy          1 ATGRLLS--------GQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 52
              |||||||        ||||||||||||||||||||||||||||||||||||||||||||
Db         22 ATGRLLSASDLDLRGGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 81

Qy         53 ESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVD 112
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 ESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVD 141

Qy        113 EPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGE 172
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 EPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGE 201

Qy        173 ETELTTPVLPEETQEEDAKKTFKESREAALNLAY 206
              |||||||||||||||||||||||||||||||||
```

```
Db        202 ETELTTPVLPEETQEEDAKKTFKESREAALNLAY 235


RESULT 49
ABO02088
ID   ABO02088 standard; protein; 382 AA.
XX
AC   ABO02088;
XX
DT   09-AUG-2003  (first entry)
XX
DE   Human secreted/transmembrane protein (PRO) #10.
XX
KW   Human; secreted and transmembrane protein; PRO; TNF-alpha;
KW   tumour necrosis factor alpha; chondrocyte cell; tumour; gene therapy;
KW   tissue typing; adrenal tumour; lung tumour; colon tumour; breast tumour;
KW   prostate tumour; rectal tumour; cervical tumour; liver tumour.
XX
OS   Homo sapiens.
XX
PN   US2003040054-A1.
XX
PD   27-FEB-2003.
XX
PF   20-JUN-2002; 2002US-00176479.
XX
PR   18-SEP-1997;   97US-0059263P.
PR   18-SEP-1997;   97US-0059266P.
PR   17-OCT-1997;   97US-0062250P.
PR   21-OCT-1997;   97US-0063486P.
PR   24-OCT-1997;   97US-0063120P.
PR   24-OCT-1997;   97US-0063121P.
PR   28-OCT-1997;   97US-0063540P.
PR   28-OCT-1997;   97US-0063541P.
PR   28-OCT-1997;   97US-0063544P.
PR   28-OCT-1997;   97US-0063564P.
PR   29-OCT-1997;   97US-0063734P.
PR   31-OCT-1997;   97US-0063870P.
PR   31-OCT-1997;   97US-0064103P.
PR   13-NOV-1997;   97US-0065311P.
PR   21-NOV-1997;   97US-0066120P.
PR   24-NOV-1997;   97US-0066466P.
PR   24-NOV-1997;   97US-0066772P.
PR   11-DEC-1997;   97US-0069335P.
PR   12-DEC-1997;   97US-0069425P.
PR   17-DEC-1997;   97US-0069870P.
PR   18-DEC-1997;   97US-0068017P.
PR   10-MAR-1998;   98US-0077450P.
PR   11-MAR-1998;   98US-0077632P.
PR   11-MAR-1998;   98US-0077649P.
PR   20-MAR-1998;   98US-0078886P.
PR   20-MAR-1998;   98US-0078939P.
PR   27-MAR-1998;   98US-0079664P.
PR   27-MAR-1998;   98US-0079786P.
PR   31-MAR-1998;   98US-0080107P.
PR   31-MAR-1998;   98US-0080194P.
PR   01-APR-1998;   98US-0080327P.
PR   01-APR-1998;   98US-0080333P.
PR   08-APR-1998;   98US-0081049P.
PR   08-APR-1998;   98US-0081070P.
PR   09-APR-1998;   98US-0081195P.
PR   15-APR-1998;   98US-0081838P.
PR   21-APR-1998;   98US-0082568P.
PR   21-APR-1998;   98US-0082569P.
PR   22-APR-1998;   98US-0082704P.
PR   22-APR-1998;   98US-0082797P.
PR   28-APR-1998;   98US-0083322P.
PR   29-APR-1998;   98US-0083495P.
PR   29-APR-1998;   98US-0083496P.
PR   29-APR-1998;   98US-0083499P.
PR   29-APR-1998;   98US-0083559P.
PR   05-MAY-1998;   98US-0084366P.
PR   06-MAY-1998;   98US-0084414P.
PR   07-MAY-1998;   98US-0084639P.
PR   07-MAY-1998;   98US-0084640P.
PR   07-MAY-1998;   98US-0084643P.
PR   15-MAY-1998;   98US-0085579P.
PR   15-MAY-1998;   98US-0085580P.
PR   15-MAY-1998;   98US-0085582P.
PR   15-MAY-1998;   98US-0085700P.
PR   18-MAY-1998;   98US-0086023P.
PR   22-MAY-1998;   98US-0086392P.
PR   22-MAY-1998;   98US-0086486P.
PR   28-MAY-1998;   98US-0087098P.
PR   28-MAY-1998;   98US-0087208P.
PR   02-JUN-1998;   98US-0087609P.
PR   02-JUN-1998;   98US-0087759P.
PR   03-JUN-1998;   98US-0087827P.
PR   04-JUN-1998;   98US-0088025P.
PR   04-JUN-1998;   98US-0088028P.
PR   04-JUN-1998;   98US-0088029P.
PR   04-JUN-1998;   98US-0088033P.
PR   04-JUN-1998;   98US-0088326P.
PR   05-JUN-1998;   98US-0088167P.
PR   05-JUN-1998;   98US-0088202P.
PR   05-JUN-1998;   98US-0088212P.
PR   05-JUN-1998;   98US-0088217P.
PR   09-JUN-1998;   98US-0088555P.
PR   10-JUN-1998;   98US-0088722P.
PR   10-JUN-1998;   98US-0088738P.
PR   10-JUN-1998;   98US-0088740P.
PR   10-JUN-1998;   98US-0088811P.
PR   10-JUN-1998;   98US-0088824P.
PR   10-JUN-1998;   98US-0088825P.
PR   10-JUN-1998;   98US-0088826P.
PR   11-JUN-1998;   98US-0088861P.
PR   11-JUN-1998;   98US-0088863P.
PR   11-JUN-1998;   98US-0088876P.
PR   12-JUN-1998;   98US-0089090P.
PR   12-JUN-1998;   98US-0089105P.
PR   16-JUN-1998;   98US-0089512P.
PR   16-JUN-1998;   98US-0089514P.
PR   17-JUN-1998;   98US-0089538P.
PR   17-JUN-1998;   98US-0089598P.
PR   17-JUN-1998;   98US-0089653P.
PR   18-JUN-1998;   98US-0089908P.
PR   19-JUN-1998;   98US-0089952P.
PR   22-JUN-1998;   98US-0090246P.
PR   22-JUN-1998;   98US-0090252P.
PR   22-JUN-1998;   98US-0090254P.
PR   24-JUN-1998;   98US-0090429P.
PR   24-JUN-1998;   98US-0090435P.
PR   24-JUN-1998;   98US-0090444P.
PR   24-JUN-1998;   98US-0090461P.
PR   24-JUN-1998;   98US-0090535P.
PR   24-JUN-1998;   98US-0090540P.
PR   25-JUN-1998;   98US-0090676P.
PR   25-JUN-1998;   98US-0090678P.
PR   25-JUN-1998;   98US-0090688P.
PR   25-JUN-1998;   98US-0090690P.
PR   25-JUN-1998;   98US-0090694P.
PR   25-JUN-1998;   98US-0090695P.
PR   25-JUN-1998;   98US-0090696P.
PR   26-JUN-1998;   98US-0090862P.
PR   26-JUN-1998;   98US-0090863P.
PR   26-JUN-1998;   98US-0091010P.
PR   01-JUL-1998;   98US-0091359P.
PR   01-JUL-1998;   98US-0091544P.
PR   02-JUL-1998;   98US-0091478P.
PR   02-JUL-1998;   98US-0091486P.
PR   02-JUL-1998;   98US-0091626P.
PR   02-JUL-1998;   98US-0091628P.
PR   02-JUL-1998;   98US-0091632P.
PR   24-JUL-1998;   98US-0094006P.
PR   04-AUG-1998;   98US-0095282P.
PR   10-AUG-1998;   98US-0095998P.
```

```
PR   10-AUG-1998;   98US-0096012P.
PR   17-AUG-1998;   98US-0096757P.
PR   17-AUG-1998;   98US-0096766P.
PR   17-AUG-1998;   98US-0096867P.
PR   17-AUG-1998;   98US-0096891P.
PR   17-AUG-1998;   98US-0096897P.
PR   18-AUG-1998;   98US-0096949P.
PR   18-AUG-1998;   98US-0096959P.
PR   18-AUG-1998;   98US-0097022P.
PR   26-AUG-1998;   98US-0097952P.
PR   26-AUG-1998;   98US-0097954P.
PR   26-AUG-1998;   98US-0097955P.
PR   26-AUG-1998;   98US-0097971P.
PR   26-AUG-1998;   98US-0097974P.
PR   26-AUG-1998;   98US-0098014P.
PR   01-SEP-1998;   98US-0098716P.
PR   01-SEP-1998;   98US-0098723P.
PR   02-SEP-1998;   98US-0098803P.
PR   02-SEP-1998;   98US-0098821P.
PR   02-SEP-1998;   98US-0098843P.
PR   09-SEP-1998;   98US-0099602P.
PR   10-SEP-1998;   98US-0099741P.
PR   10-SEP-1998;   98US-0099754P.
PR   10-SEP-1998;   98US-0099763P.
PR   10-SEP-1998;   98US-0099812P.
PR   15-SEP-1998;   98US-0100388P.
PR   16-SEP-1998;   98US-0100662P.
PR   16-SEP-1998;   98US-0100664P.
PR   16-SEP-1998;   98US-0101751P.
PR   16-SEP-1998;   98WO-US019330.
PR   17-SEP-1998;   98US-0100683P.
PR   17-SEP-1998;   98US-0100684P.
PR   17-SEP-1998;   98US-0100919P.
PR   17-SEP-1998;   98US-0100930P.
PR   18-SEP-1998;   98US-0100849P.
PR   18-SEP-1998;   98US-0101014P.
PR   18-SEP-1998;   98US-0101068P.
PR   23-SEP-1998;   98US-0101471P.
PR   23-SEP-1998;   98US-0101472P.
PR   23-SEP-1998;   98US-0101475P.
PR   23-SEP-1998;   98US-0101477P.
PR   24-SEP-1998;   98US-0101738P.
PR   24-SEP-1998;   98US-0101739P.
PR   24-SEP-1998;   98US-0101743P.
PR   24-SEP-1998;   98US-0101922P.
PR   25-SEP-1998;   98US-0101786P.
PR   29-SEP-1998;   98US-0102207P.
PR   29-SEP-1998;   98US-0102240P.
PR   29-SEP-1998;   98US-0102330P.
PR   29-SEP-1998;   98US-0102331P.
PR   30-SEP-1998;   98US-0102487P.
PR   30-SEP-1998;   98US-0102570P.
PR   30-SEP-1998;   98US-0102571P.
PR   01-OCT-1998;   98US-0102684P.
PR   01-OCT-1998;   98US-0102687P.
PR   02-OCT-1998;   98US-0102965P.
PR   06-OCT-1998;   98US-0103258P.
PR   06-OCT-1998;   98US-0103449P.
PR   07-OCT-1998;   98US-0103395P.

  Query Match             98.7%;  Score 1101;  DB 6;  Length 382;
  Best Local Similarity   96.3%;  Pred. No. 5.2e-103;
  Matches  206;  Conservative    0;  Mismatches    0;  Indels    8;  Gaps    1;

Qy          1 ATGRLLS--------GQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI  52
              |||||||        ||||||||||||||||||||||||||||||||||||||||||||
Db         22 ATGRLLSASDLDLRGGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI  81

Qy         53 ESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVD  112
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 ESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVD  141

Qy        113 EPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGE  172
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 EPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGE  201

Qy        173 ETELTTPVLPEETQEEDAKKTFKESREAALNLAY  206
              |||||||||||||||||||||||||||||||||
Db        202 ETELTTPVLPEETQEEDAKKTFKESREAALNLAY  235


RESULT 50
ABU54370
ID   ABU54370 standard; protein; 382 AA.
XX
AC   ABU54370;
XX
DT   10-MAR-2003  (first entry)
XX
DE   Human secreted/transmembrane protein PRO234.
XX
KW   Human; PRO; secreted protein; transmembrane protein; enterocolitis;
KW   gastrointestinal ulceration; skin disease;
KW   abnormal keratinocyte differentiation; psoriasis; epithelial cancer;
KW   squamous cell carcinoma; Alzheimer's disease; Parkinson's disease;
KW   amyotrophic lateral sclerosis; inflammatory disease;
KW   rheumatoid arthritis; asthma; multiple sclerosis; organ failure;
KW   atherosclerosis; cardiac injury; infertility; birth defect;
KW   premature aging; AIDS; acquired immunodeficiency syndrome; cancer;
KW   diabetic complication; wound repair.
XX
OS   Homo sapiens.
XX
PN   US2002132240-A1.
XX
PD   19-SEP-2002.
XX
PF   18-JUL-2001; 2001US-00909320.
XX
PR   17-SEP-1997;   97US-0059113P.
PR   17-SEP-1997;   97US-0059115P.
PR   17-SEP-1997;   97US-0059117P.
PR   17-SEP-1997;   97US-0059119P.
PR   17-SEP-1997;   97US-0059121P.
PR   17-SEP-1997;   97US-0059122P.
PR   17-SEP-1997;   97US-0059184P.
PR   18-SEP-1997;   97US-0059263P.
PR   18-SEP-1997;   97US-0059266P.
PR   15-OCT-1997;   97US-0062125P.
PR   17-OCT-1997;   97US-0062285P.
PR   17-OCT-1997;   97US-0062287P.
PR   21-OCT-1997;   97US-0063486P.
PR   24-OCT-1997;   97US-0062814P.
PR   24-OCT-1997;   97US-0062816P.
PR   24-OCT-1997;   97US-0063045P.
PR   24-OCT-1997;   97US-0063120P.
PR   24-OCT-1997;   97US-0063121P.
PR   24-OCT-1997;   97US-0063127P.
PR   24-OCT-1997;   97US-0063128P.
PR   27-OCT-1997;   97US-0063327P.
PR   27-OCT-1997;   97US-0063329P.
PR   28-OCT-1997;   97US-0063541P.
PR   28-OCT-1997;   97US-0063542P.
PR   28-OCT-1997;   97US-0063544P.
PR   28-OCT-1997;   97US-0063549P.
PR   28-OCT-1997;   97US-0063550P.
PR   28-OCT-1997;   97US-0063564P.
PR   29-OCT-1997;   97US-0063435P.
PR   29-OCT-1997;   97US-0063704P.
PR   29-OCT-1997;   97US-0063732P.
PR   29-OCT-1997;   97US-0063734P.
PR   29-OCT-1997;   97US-0063735P.
PR   29-OCT-1997;   97US-0063738P.
PR   29-OCT-1997;   97US-0064215P.
PR   31-OCT-1997;   97US-0063870P.
PR   31-OCT-1997;   97US-0064103P.
```

PR 03-NOV-1997; 97US-0064248P.
PR 07-NOV-1997; 97US-0064809P.
PR 12-NOV-1997; 97US-0065186P.
PR 17-NOV-1997; 97US-0065846P.
PR 18-NOV-1997; 97US-0065693P.
PR 21-NOV-1997; 97US-0066120P.
PR 21-NOV-1997; 97US-0066364P.
PR 24-NOV-1997; 97US-0066453P.
PR 24-NOV-1997; 97US-0066466P.
PR 24-NOV-1997; 97US-0066511P.
PR 24-NOV-1997; 97US-0066770P.
PR 24-NOV-1997; 97US-0066772P.
PR 10-SEP-1998; 98WO-US018824.
PR 14-SEP-1998; 98WO-US019177.
PR 16-SEP-1998; 98WO-US019330.
PR 17-SEP-1998; 98WO-US019437.
PR 01-DEC-1998; 98WO-US025108.
PR 08-SEP-1999; 99WO-US020594.
PR 13-SEP-1999; 99WO-US020944.
PR 15-SEP-1999; 99WO-US021090.
PR 15-SEP-1999; 99WO-US021547.
PR 05-OCT-1999; 99WO-US023089.
PR 29-NOV-1999; 99WO-US028214.
PR 30-NOV-1999; 99WO-US028313.
PR 01-DEC-1999; 99WO-US028301.
PR 02-DEC-1999; 99WO-US028564.
PR 02-DEC-1999; 99WO-US028565.
PR 16-DEC-1999; 99WO-US030095.
PR 20-DEC-1999; 99WO-US030911.
PR 20-DEC-1999; 99WO-US030999.
PR 06-JAN-2000; 2000WO-US000219.
PR 11-FEB-2000; 2000WO-US003565.
PR 22-FEB-2000; 2000WO-US004414.
PR 24-FEB-2000; 2000WO-US005004.
PR 02-MAR-2000; 2000WO-US005841.
PR 20-MAR-2000; 2000WO-US007377.
PR 30-MAR-2000; 2000WO-US008439.
PR 22-MAY-2000; 2000WO-US014042.
PR 02-JUN-2000; 2000WO-US015264.
PR 28-JUL-2000; 2000WO-US020710.
PR 24-AUG-2000; 2000WO-US023328.
PR 18-SEP-2000; 2000US-00665350.
XX
PA (GETH) GENENTECH INC.
XX
PI Ashkenazi A, Botstein D, Desnoyers L, Eaton DL, Ferrara N;
PI Filvaroff E, Fong S, Gao W, Gerber H, Gerritsen ME, Goddard A;
PI Godowski PJ, Grimaldi JC, Gurney AL, Hillan KJ, Kljavin IJ;
PI Mather JP, Pan J, Paoni NF, Roy MA, Stewart TA, Tumas D;
PI Williams PM, Wood WI;
XX
DR WPI; 2003-147434/14.
DR N-PSDB; ABX71560.
XX
PT New PRO polypeptides and nucleic acid molecules, useful in diagnosing or
PT treating inflammatory diseases, organ failure, atherosclerosis, cardiac
PT injury, infertility, cancer, AIDS, Alzheimer's disease or Parkinson's
PT disease.
XX
PS Claim 12; Fig 50; 473pp; English.
XX
CC The invention relates to an isolated PRO polypeptide having at least 80%
CC amino acid sequence identity to: (a) any one of 61 fully defined amino
CC acid sequences given in the specification (appearing as ABU54347-
CC ABU54407); (b) an amino acid sequence encoded by the nucleotide sequence
CC deposited under American Type Culture Collection (accession numbers
CC listed in the specification); (c) any one of the PRO sequences which
CC lacks its associated signal peptide; (d) an extracellular domain of the
CC PRO polypeptide with its associated signal peptide; or (e) an
CC extracellular domain of the PRO polypeptide which lacks its associated
CC signal peptide. Also include are the nucleic acids encoding the PRO
CC polypeptides, vectors, host cells and anti-PRO antibodies. The PRO
CC polypeptides and nucleic acids are useful in diagnosing or treating
CC enterocolitis, gastrointestinal ulceration, skin diseases associated with
CC abnormal keratinocyte differentiation, e.g. psoriasis or epithelial
CC cancers such as squamous cell carcinoma, Alzheimer's disease, Parkinson's
CC disease, amyotrophic lateral sclerosis, inflammatory diseases, e.g.
CC rheumatoid arthritis, asthma or multiple sclerosis, organ failure,
CC atherosclerosis, cardiac injury, infertility, birth defects, premature
CC aging, AIDS, cancer, diabetic complications, or mutations in general. The
CC polypeptides are also useful for wound repair and associated therapies
CC concerned with re-growth of tissue. The nucleotide sequences may be used
CC as hybridisation probes in chromosome and gene mapping, or in generating
CC antisense RNA and DNA. PRO nucleic acids are also useful in preparing PRO
CC polypeptides, in assays to identify other proteins or molecules involved
CC in binding reaction, to generate transgenic animals or knockout animals,
CC which in turn are useful in the development and screening of
CC therapeutically useful reagents, for chromosome identification, and
CC tissue typing. The PRO polypeptides and nucleic acid molecules are also
CC useful in gene therapy, and as molecular weight markers for protein
CC electrophoresis purposes. The anti-PRO antibodies may be used in
CC diagnostic assays for PRO, or for the affinity purification of PRO from
CC recombinant cell culture or natural sources. The present sequence
CC represents a PRO polypeptide
XX
SQ Sequence 382 AA;

Query Match 98.7%; Score 1101; DB 6; Length 382;
Best Local Similarity 96.3%; Pred. No. 5.2e-103;
Matches 206; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

```
Qy          1 ATGRLLS--------GQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 52
              |||||||        ||||||||||||||||||||||||||||||||||||||||||||
Db         22 ATGRLLSASDLDLRGGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 81

Qy         53 ESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVD 112
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 ESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVD 141

Qy        113 EPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGE 172
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 EPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGE 201

Qy        173 ETELTTPVLPEETQEEDAKKTFKESREAALNLAY 206
              |||||||||||||||||||||||||||||||||
Db        202 ETELTTPVLPEETQEEDAKKTFKESREAALNLAY 235
```

Search completed: September 9, 2004, 22:36:20
Job time : 127 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: September 9, 2004, 22:18:00 ; Search time 60 Seconds
(without alignments)
1761.213 Million cell updates/sec

Title: US-09-887-855-2
Perfect score: 2000
Sequence: 1 MRPGTALQAVLLAVLLVGLR.....PDQMGRSKESGWVENEIYGY 374

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2000	100.0	374	3	AAY93948	Aay93948 Amino aci
2	2000	100.0	374	4	AAE03651	Aae03651 Human ext
3	2000	100.0	374	5	ABB90203	Abb90203 Human pol
4	2000	100.0	374	6	ADA54522	Ada54522 Human pro
5	1996	99.8	374	6	ADA54574	Ada54574 Human pro
6	1990	99.5	374	3	AAY91490	Aay91490 Human sec
7	1986	99.3	382	2	AAY13367	Aay13367 Amino aci
8	1986	99.3	382	3	ADC78457	Adc78457 Human PRO
9	1986	99.3	382	4	AAB80235	Aab80235 Human PRO
10	1986	99.3	382	4	AAU29033	Aau29033 Human PRO
11	1986	99.3	382	6	ABU58409	Abu58409 Human PRO
12	1986	99.3	382	6	ABU71613	Abu71613 Human PRO
13	1986	99.3	382	6	ABU87957	Abu87957 Novel hum
14	1986	99.3	382	6	ABU84272	Abu84272 Human sec
15	1986	99.3	382	6	ABR66146	Abr66146 Human sec
16	1986	99.3	382	6	ABR65536	Abr65536 Human sec
17	1986	99.3	382	6	ABU99476	Abu99476 Human sec
18	1986	99.3	382	6	ABU82715	Abu82715 Human PRO
19	1986	99.3	382	6	ABU89836	Abu89836 Novel hum
20	1986	99.3	382	6	ABU71468	Abu71468 Human PRO
21	1986	99.3	382	6	ABR68085	Abr68085 Human sec
22	1986	99.3	382	6	ABU96138	Abu96138 Novel hum
23	1986	99.3	382	6	ABU92569	Abu92569 Human sec
24	1986	99.3	382	6	ABO08646	Abo08646 Human sec
25	1986	99.3	382	6	ABO02698	Abo02698 Human sec
26	1986	99.3	382	6	ABR74852	Abr74852 Human sec
27	1986	99.3	382	6	ABR94614	Abr94614 Human sec
28	1986	99.3	382	6	ABU85587	Abu85587 Human PRO
29	1986	99.3	382	6	ABU98747	Abu98747 Novel hum
30	1986	99.3	382	6	ABU97962	Abu97962 Novel hum
31	1986	99.3	382	6	ABU91668	Abu91668 Novel hum
32	1986	99.3	382	6	ABU71914	Abu71914 Human sec
33	1986	99.3	382	6	ABU89361	Abu89361 Human PRO
34	1986	99.3	382	6	ABU86202	Abu86202 Human sec
35	1986	99.3	382	6	ABU67415	Abu67415 Human sec
36	1986	99.3	382	6	ABU80443	Abu80443 Human PRO
37	1986	99.3	382	6	ABO01797	Abo01797 Novel hum
38	1986	99.3	382	6	ABR99361	Abr99361 Human sec
39	1986	99.3	382	6	ABR98751	Abr98751 Human sec
40	1986	99.3	382	6	ABO16274	Abo16274 Human sec
41	1986	99.3	382	6	ABR92174	Abr92174 Human sec
42	1986	99.3	382	6	ABO18815	Abo18815 Human sec
43	1986	99.3	382	6	ABR78236	Abr78236 Human sec
44	1986	99.3	382	6	ABU84972	Abu84972 Novel hum
45	1986	99.3	382	6	ABO00111	Abo00111 Novel hum

ALIGNMENTS

RESULT 1
AAY93948
ID AAY93948 standard; protein; 374 AA.
XX
AC AAY93948;
XX
DT 03-OCT-2000 (first entry)
XX
DE Amino acid sequence of a lectin ss3939 polypeptide.
XX
KW Human; lectin ss3939; chromosome 11; gene therapy.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Peptide 1. .21
FT /note= "signal peptide"
FT Domain 22. .227
FT /note= "extracellular coding region"
FT Domain 228. .248
FT /note= "predicted transmembrane domain"
FT Domain 249. .374
FT /note= "predicted cytoplasmic or intracellular domain"
XX
PN WO200039296-A1.
XX
PD 06-JUL-2000.
XX
PF 22-DEC-1999; 99WO-US030523.
XX
PR 23-DEC-1998; 98US-0113820P.
XX
PA (IMMV) IMMUNEX CORP.
XX
PI Anderson DA;
XX
DR WPI; 2000-452394/39.
DR N-PSDB; AAA57382.
XX
PT ss3939 nucleic acids, polypeptides and antibodies, useful for identifying
PT human chromosome 11 and diseases associated with it.
XX
PS Claim 12; Page 8; 73pp; English.
XX
CC The present sequence represents a human lectin ss3939 polypeptide. The
CC polynucleotide sequence is a source of probes, which may be used to
CC identify nucleic acids encoding ss3939 proteins, to identify human
CC chromosome number 11, to map genes on human chromosome number 11, to

```
CC   identify diseases associated with chromosome 11, as single-stranded sense
CC   or antisense oligonucleotides to inhibit expression of polypeptides
CC   encoded by the ss3939 gene, and for gene therapy. The ss3939 polypeptides
CC   may be useful for developing treatments for diseases (none specified)
CC   associated with defective or insufficient amounts of the polypeptides.
CC   The antibodies may be useful for detecting the presence of ss3939
CC   polypeptides
XX
SQ   Sequence 374 AA;

  Query Match             100.0%;  Score 2000;  DB 3;  Length 374;
  Best Local Similarity   100.0%;  Pred. No. 2.1e-182;
  Matches  374;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MRPGTALQAVLLAVLLVGLRAATGRLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAK 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MRPGTALQAVLLAVLLVGLRAATGRLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAK 60

Qy         61 EACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 EACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWT 120

Qy        121 DGSISQFRNWYVDEPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSD 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 DGSISQFRNWYVDEPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSD 180

Qy        181 EKPAVPSREAEGEETELTTPVLPEETQEEDAKKTFKESREAALNLAYILIPSIPLLLLLV 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 EKPAVPSREAEGEETELTTPVLPEETQEEDAKKTFKESREAALNLAYILIPSIPLLLLLV 240

Qy        241 VTTVVCWVWICRKRKREQPDPSTKKQHTIWPSPHQGNSPDLEVYNVIRKQSEADLAETRP 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 VTTVVCWVWICRKRKREQPDPSTKKQHTIWPSPHQGNSPDLEVYNVIRKQSEADLAETRP 300

Qy        301 DLKNISFRVCSGEATPDDMSCDYDNMAVNPSESGFVTLVSVESGFVTNDIYEFSPDQMGR 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 DLKNISFRVCSGEATPDDMSCDYDNMAVNPSESGFVTLVSVESGFVTNDIYEFSPDQMGR 360

Qy        361 SKESGWVENEIYGY 374
              ||||||||||||||
Db        361 SKESGWVENEIYGY 374


RESULT 2
AAE03651
ID   AAE03651 standard; protein; 374 AA.
XX
AC   AAE03651;
XX
DT   06-AUG-2001  (first entry)
XX
DE   Human extracellular matrix and cell adhesion molecule-15 (XMAD-15).
XX
KW   Human; extracellular matrix and cell adhesion molecule; XMAD;
KW   gene therapy; genetic disorder; adrenoleukodystrophy; leukaemia;
KW   Down's syndrome; cystic fibrosis; Gaucher's disease; myotonic dystrophy;
KW   sickle cell anaemia; thalassaemia; autoimmune disorder; adenocarcinoma;
KW   inflammatory disorder; acquired immune deficiency syndrome; AIDS;
KW   Addison's disease; allergy; anaemia; asthma; atherosclerosis; melanoma;
KW   Crohn's disease; diabetes mellitus; atopic dermatitis; lymphoma; cancer;
KW   glomerulonephritis; multiple sclerosis; Grave's disease; osteoarthritis;
KW   osteoporosis; psoriasis; rheumatoid arthritis; ulcerative colitis;
KW   infection; cell proliferative disorder; actinic keratosis; myeloma;
KW   arteriosclerosis; nootropic; anticonvulsant; antithyroid; nephrotropic;
KW   neuroprotective; dermatological.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Peptide         1. .24
FT                   /label= Signal_peptide
FT   Protein         25. .374
FT                   /note= "Mature human extracellular matrix and cell
FT                   adhesion molecule (XMAD)"
FT   Domain          46. .63
FT                   /note= "C-type lectin domain"
FT   Domain          163. .176
FT                   /note= "C-type lectin domain"
FT   Domain          224. .247
FT                   /note= "Transmembrane motif"
FT   Domain          328. .348
FT                   /note= "Transmembrane motif"
XX
PN   WO200142285-A2.
XX
PD   14-JUN-2001.
XX
PF   05-DEC-2000; 2000WO-US032990.
XX
PR   10-DEC-1999;   99US-0172852P.
PR   16-DEC-1999;   99US-0172354P.
XX
PA   (INCY-) INCYTE GENOMICS INC.
XX
PI   Yue H,  Tang YT,  Lal P,  Burford N,  Azimzai Y,  Patterson C;
PI   Baughn MR,  Lu DAM,  Shah P,  Au-Young J;
XX
DR   WPI; 2001-381632/40.
DR   N-PSDB; AAD08059.
XX
PT   New human extracellular matrix and cell adhesion molecules and
PT   polynucleotide sequences encoding them, useful for diagnosis, prevention,
PT   treatment of genetic, autoimmune and cell proliferative disorders.
XX
PS   Claim 1; Page 108-109; 135pp; English.
XX
CC   The present sequence is a human extracellular matrix and cell adhesion
CC   molecule (XMAD). The XMAD is used for screening a compound for
CC   effectiveness as an agonist or antagonist of XMAD. The identified agonist
CC   or antagoinst are used for treating a disease or condition associated
CC   with decreased or increased expression of functional XMAD. The
CC   polynucleotides encoding XMAD are useful in somatic or germline gene
CC   therapy to correct a genetic deficiency, to express a conditionally
CC   lethal gene product and to express a protein which affords protection
CC   against intracellular parasites and also for diagnosis of disorders
CC   associated with expression of XMAD. They are also used for generating
CC   hybridisation probes useful in mapping the naturally occurring genomic
CC   sequences and to create knock in humanised animals (pigs) or transgenic
CC   animals (mice or rats) to model human diseases. Oligonucleotide or longer
CC   fragments derived from the polynucleotide sequences may be used as
CC   elements on a microarray. Antibodies which specifically bind XMAD may be
CC   used for the diagnosis of disorders associated with the expression of
CC   XMAD, or in assays to monitor patients being treated with XMAD. Diseases
CC   diagnosed, prevented or treated include genetic disorders such as
CC   adrenoleukodystrophy, Down's syndrome, cystic fibrosis, Gaucher's
CC   disease, myotonic dystrophy, sickle cell anaemia, thalassaemia,
CC   autoimmune/inflammatory disorders such as acquired immune deficiency
CC   syndrome (AIDS), Addison's disease, allergies, anaemia, asthma,
CC   atherosclerosis, Crohn's disease, diabetes mellitus, atopic dermatitis,
CC   glomerulonephritis, multiple sclerosis, Grave's disease, osteoarthritis,
CC   osteoporosis, psoriasis, rheumatoid arthritis, ulcerative colitis,
CC   bacterial, fungal, parasitic, protozoal and helminthic infections and
CC   cell proliferative disorders such as actinic keratosis, arteriosclerosis
CC   and cancer including breast, bladder, bone marrow, brain and uterus
CC   cancer, leukaemia, adenocarcinoma, lymphoma, melanoma and myeloma
XX
SQ   Sequence 374 AA;

  Query Match             100.0%;  Score 2000;  DB 4;  Length 374;
  Best Local Similarity   100.0%;  Pred. No. 2.1e-182;
  Matches  374;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MRPGTALQAVLLAVLLVGLRAATGRLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAK 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MRPGTALQAVLLAVLLVGLRAATGRLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAK 60
```

```
Qy         61 EACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 EACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWT 120

Qy        121 DGSISQFRNWYVDEPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSD 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 DGSISQFRNWYVDEPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSD 180

Qy        181 EKPAVPSREAEGEETELTTPVLPEETQEEDAKKTFKESREAALNLAYILIPSIPLLLLLV 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 EKPAVPSREAEGEETELTTPVLPEETQEEDAKKTFKESREAALNLAYILIPSIPLLLLLV 240

Qy        241 VTTVVCWVWICRKRKREQPDPSTKKQHTIWPSPHQGNSPDLEVYNVIRKQSEADLAETRP 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 VTTVVCWVWICRKRKREQPDPSTKKQHTIWPSPHQGNSPDLEVYNVIRKQSEADLAETRP 300

Qy        301 DLKNISFRVCSGEATPDDMSCDYDNMAVNPSESGFVTLVSVESGFVTNDIYEFSPDQMGR 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 DLKNISFRVCSGEATPDDMSCDYDNMAVNPSESGFVTLVSVESGFVTNDIYEFSPDQMGR 360

Qy        361 SKESGWVENEIYGY 374
              ||||||||||||||
Db        361 SKESGWVENEIYGY 374
```

RESULT 3
ABB90203
ID ABB90203 standard; protein; 374 AA.
XX
AC ABB90203;
XX
DT 24-MAY-2002 (first entry)
XX
DE Human polypeptide SEQ ID NO 2579.
XX
KW Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
KW antiallergic; hepatotropic; antidiabetic; antiinflammatory; antiulcer;
KW vulnerary; anticonvulsant; antibacterial; antifungal; antiparasitic;
KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
KW neurological disease; infection; human; secreted protein.
XX
OS Homo sapiens.
XX
PN WO200190304-A2.
XX
PD 29-NOV-2001.
XX
PF 18-MAY-2001; 2001WO-US016450.
XX
PR 19-MAY-2000; 2000US-0205515P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Birse CE, Rosen CA;
XX
DR WPI; 2002-122018/16.
DR N-PSDB; ABL90612.
XX
PT Novel 1405 isolated polypeptides, useful for diagnosis, treatment and
PT prevention of neural, immune system, muscular, reproductive,
PT gastrointestinal, pulmonary, cardiovascular, renal and proliferative
PT disorders.
XX
PS Claim 11; SEQ ID NO 2579; 2081pp + Sequence Listing; English.
XX
CC The invention relates to novel genes (ABL89449-ABL90853) and proteins
CC (ABB89040-ABB90444) useful for preventing, treating or ameliorating
CC medical conditions e.g. by protein or gene therapy. The genes are
CC isolated from a range of human tissues disclosed in the specification.
CC The nucleic acids, proteins, antibodies and (ant)agonists are useful in
CC the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and
CC ovarian cancer and other cancers of the adrenal gland, bone, bone marrow,
CC breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune
CC disorders e.g. Addison's disease, allergies, autoimmune haemolytic
CC anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,
CC multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c)
CC cardiovascular disorders such as myocardial ischaemias; (d) wound healing
CC ; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f)
CC infectious diseases such as viral, bacterial, fungal and parasitic
CC infections. Note: The sequence data for this patent did not form part of
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 374 AA;

```
  Query Match           100.0%;  Score 2000;  DB 5;  Length 374;
  Best Local Similarity  100.0%;  Pred. No. 2.1e-182;
  Matches  374;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MRPGTALQAVLLAVLLVGLRAATGRLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAK 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MRPGTALQAVLLAVLLVGLRAATGRLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAK 60

Qy         61 EACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 EACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWT 120

Qy        121 DGSISQFRNWYVDEPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSD 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 DGSISQFRNWYVDEPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSD 180

Qy        181 EKPAVPSREAEGEETELTTPVLPEETQEEDAKKTFKESREAALNLAYILIPSIPLLLLLV 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 EKPAVPSREAEGEETELTTPVLPEETQEEDAKKTFKESREAALNLAYILIPSIPLLLLLV 240

Qy        241 VTTVVCWVWICRKRKREQPDPSTKKQHTIWPSPHQGNSPDLEVYNVIRKQSEADLAETRP 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 VTTVVCWVWICRKRKREQPDPSTKKQHTIWPSPHQGNSPDLEVYNVIRKQSEADLAETRP 300

Qy        301 DLKNISFRVCSGEATPDDMSCDYDNMAVNPSESGFVTLVSVESGFVTNDIYEFSPDQMGR 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 DLKNISFRVCSGEATPDDMSCDYDNMAVNPSESGFVTLVSVESGFVTNDIYEFSPDQMGR 360

Qy        361 SKESGWVENEIYGY 374
              ||||||||||||||
Db        361 SKESGWVENEIYGY 374
```

RESULT 4
ADA54522
ID ADA54522 standard; protein; 374 AA.
XX
AC ADA54522;
XX
DT 20-NOV-2003 (first entry)
XX
DE Human protein, SEQ ID 2090.
XX
KW Cytostatic; Anti-inflammatory; Osteopathic; Neuroprotective; Nootropic;
KW Gene Therapy; human; secretory protein; membrane proteins; cancer;
KW inflammatory disease; osteoporosis; neurological disease.
XX
OS Homo sapiens.
XX
PN EP1293569-A2.
XX
PD 19-MAR-2003.
XX
PF 21-MAR-2002; 2002EP-00006586.
XX
PR 14-SEP-2001; 2001JP-00328381.
PR 24-JAN-2002; 2002US-0350435P.
XX
PA (HELI-) HELIX RES INST.
PA (REAS-) RES ASSOC BIOTECHNOLOGY.
XX

PI Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;
PI Yamamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Irie R, Tamechika I;
PI Seki N, Yoshikawa T, Otsuka M, Nagahari K, Masuho Y;
XX
DR WPI; 2003-395539/38.
DR N-PSDB; ADA52883.
XX
PT New polynucleotides encoding full-length polypeptides, e.g. secretory
PT and/or membrane proteins, useful for developing medicines for diseases in
PT which the gene is involved, or as target molecules for gene therapy.
XX
PS Claim 14; SEQ ID NO 2090; 205pp; English.
XX
CC The present invention relates to novel human secretory or membrane
CC proteins (ADA54072-ADA55710) and their coding sequences (ADA52433-
CC ADA54071). The coding sequences are useful in the gene therapy of
CC diseases caused by abnormalities of the proteins, e.g. cancer,
CC inflammatory diseases, osteoporosis or neurological disease.
XX
SQ Sequence 374 AA;

Query Match 100.0%; Score 2000; DB 6; Length 374;
Best Local Similarity 100.0%; Pred. No. 2.1e-182;
Matches 374; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MRPGTALQAVLLAVLLVGLRAATGRLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAK 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MRPGTALQAVLLAVLLVGLRAATGRLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAK 60

Qy         61 EACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWT 120
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 EACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWT 120

Qy        121 DGSISQFRNWYVDEPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSD 180
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 DGSISQFRNWYVDEPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSD 180

Qy        181 EKPAVPSREAEGEETELTTPVLPEETQEEDAKKTFKESREAALNLAYILIPSIPLLLLLV 240
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 EKPAVPSREAEGEETELTTPVLPEETQEEDAKKTFKESREAALNLAYILIPSIPLLLLLV 240

Qy        241 VTTVVCWVWICRKRKREQPDPSTKKQHTIWPSPHQGNSPDLEVYNVIRKQSEADLAETRP 300
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 VTTVVCWVWICRKRKREQPDPSTKKQHTIWPSPHQGNSPDLEVYNVIRKQSEADLAETRP 300

Qy        301 DLKNISFRVCSGEATPDDMSCDYDNMAVNPSESGFVTLVSVESGFVTNDIYEFSPDQMGR 360
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 DLKNISFRVCSGEATPDDMSCDYDNMAVNPSESGFVTLVSVESGFVTNDIYEFSPDQMGR 360

Qy        361 SKESGWVENEIYGY 374
              ||||||||||||||
Db        361 SKESGWVENEIYGY 374
```

RESULT 5
ADA54574
ID ADA54574 standard; protein; 374 AA.
XX
AC ADA54574;
XX
DT 20-NOV-2003 (first entry)
XX
DE Human protein, SEQ ID 2142.
XX
KW Cytostatic; Anti-inflammatory; Osteopathic; Neuroprotective; Nootropic;
KW Gene Therapy; human; secretory protein; membrane proteins; cancer;
KW inflammatory disease; osteoporosis; neurological disease.
XX
OS Homo sapiens.
XX
PN EP1293569-A2.
XX
PD 19-MAR-2003.
XX
PF 21-MAR-2002; 2002EP-00006586.
XX
PR 14-SEP-2001; 2001JP-00328381.
PR 24-JAN-2002; 2002US-0350435P.
XX
PA (HELI-) HELIX RES INST.
PA (REAS-) RES ASSOC BIOTECHNOLOGY.
XX
PI Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;
PI Yamamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Irie R, Tamechika I;
PI Seki N, Yoshikawa T, Otsuka M, Nagahari K, Masuho Y;
XX
DR WPI; 2003-395539/38.
DR N-PSDB; ADA52935.
XX
PT New polynucleotides encoding full-length polypeptides, e.g. secretory
PT and/or membrane proteins, useful for developing medicines for diseases in
PT which the gene is involved, or as target molecules for gene therapy.
XX
PS Claim 14; SEQ ID NO 2142; 205pp; English.
XX
CC The present invention relates to novel human secretory or membrane
CC proteins (ADA54072-ADA55710) and their coding sequences (ADA52433-
CC ADA54071). The coding sequences are useful in the gene therapy of
CC diseases caused by abnormalities of the proteins, e.g. cancer,
CC inflammatory diseases, osteoporosis or neurological disease.
XX
SQ Sequence 374 AA;

Query Match 99.8%; Score 1996; DB 6; Length 374;
Best Local Similarity 99.7%; Pred. No. 5.1e-182;
Matches 373; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 MRPGTALQAVLLAVLLVGLRAATGRLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAK 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MRPGTALQAVLLAVLLVGLRAATGRLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAK 60

Qy         61 EACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWT 120
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 EACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWT 120

Qy        121 DGSISQFRNWYVDEPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSD 180
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 DGSISQFRNWYVDEPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSD 180

Qy        181 EKPAVPSREAEGEETELTTPVLPEETQEEDAKKTFKESREAALNLAYILIPSIPLLLLLV 240
              |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db        181 EKPAVPSREAEGEETELTTPVLPEETQEEDTKKTFKESREAALNLAYILIPSIPLLLLLV 240

Qy        241 VTTVVCWVWICRKRKREQPDPSTKKQHTIWPSPHQGNSPDLEVYNVIRKQSEADLAETRP 300
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 VTTVVCWVWICRKRKREQPDPSTKKQHTIWPSPHQGNSPDLEVYNVIRKQSEADLAETRP 300

Qy        301 DLKNISFRVCSGEATPDDMSCDYDNMAVNPSESGFVTLVSVESGFVTNDIYEFSPDQMGR 360
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 DLKNISFRVCSGEATPDDMSCDYDNMAVNPSESGFVTLVSVESGFVTNDIYEFSPDQMGR 360

Qy        361 SKESGWVENEIYGY 374
              ||||||||||||||
Db        361 SKESGWVENEIYGY 374
```

RESULT 6
AAY91490
ID AAY91490 standard; protein; 374 AA.
XX
AC AAY91490;
XX
DT 29-JUN-2000 (first entry)
XX
DE Human secreted protein sequence encoded by gene 40 SEQ ID NO:163.
XX

KW Human; secreted protein; diagnosis; cytostatic; immunosuppressive;
KW antiHIV; antiinflammatory; nootropic; neuroprotective; antiallergic;
KW osteopathic; antiarthritic; antibacterial; antidiabetic; antiasthma;
KW antipsoriatic; cardiant; gene therapy; cancer; neurological disorder;
KW immune disease; inflammation; blood disorder; tumour.
XX
OS Homo sapiens.
XX
PN WO200006698-A1.
XX
PD 10-FEB-2000.
XX
PF 29-JUL-1999; 99WO-US017130.
XX
PR 30-JUL-1998; 98US-0094657P.
PR 05-AUG-1998; 98US-0095486P.
PR 06-AUG-1998; 98US-0095454P.
PR 06-AUG-1998; 98US-0095455P.
PR 12-AUG-1998; 98US-0096319P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Komatsoulis GA, Rosen CA, Ruben SM, Duan R, Moore PA, Shi Y;
PI Lafleur D, Wei Y, Ni J, Florence KA, Young PE, Brewer LA;
PI Soppet DR, Endress GA, Ebner R, Olsen HS, Mucenski M;
XX
DR WPI; 2000-195282/17.
DR N-PSDB; AAA26385.
XX
PT New isolated human genes and the secreted polypeptides they encode,
PT useful for diagnosis and treatment of e.g. cancers, neurological
PT disorders, immune diseases, inflammation or blood disorders.
XX
PS Claim 11; Page 483-484; 634pp; English.
XX
CC The polynucleotide sequences given in AAA26346 to AAA26458 encode the
CC human secreted proteins given in AAY91451 to AAY91691. The human secreted
CC proteins can have activities based on the tissues and cells they are
CC expressed in. Examples of the activities are: cytostatic;
CC immunosuppressive; antiHIV; antiinflammatory; nootropic; neuroprotective;
CC antiallergic; osteopathic; antiarthritic; antibacterial; antidiabetic;
CC antiasthma; antipsoriatic; and cardiant. The polynucleotides and their
CC corresponding secreted proteins are useful for preventing, treating or
CC ameliorating medical conditions, e.g. by protein or gene therapy. Also
CC pathological conditions can be diagnosed by determining the amount of the
CC proteins in a sample or by determining the presence of mutations in the
CC polynucleotides. Specific uses are described for each of the
CC polynucleotides, based on which tissues they are most highly expressed
CC in, and include developing products for the diagnosis or treatment of
CC cancer, tumours, neurodegenerative disorders, developmental abnormalities
CC and foetal deficiencies, blood disorders, diseases of the immune system,
CC autoimmune diseases, hepatic and renal disease, inflammation, allergies,
CC Alzheimer's and behavioural disorders, schizophrenia, osteoporosis,
CC arthritis, infections, AIDS, spinal cord injuries, transplant rejection,
CC diabetes, asthma, sepsis, acne, psoriasis, cardiovascular disorders,
CC reproductive disorders, gastrointestinal disorders, respiratory disorders
CC and metabolic disorders. The proteins or polynucleotides can also be used
CC as food additives or preservatives. The proteins are also useful for
CC identifying their binding partners. AAA26337 to AAA26345 and AAY91450 are
CC sequences used in the exemplification of the present invention
XX
SQ Sequence 374 AA;

Query Match 99.5%; Score 1990; DB 3; Length 374;
Best Local Similarity 99.5%; Pred. No. 1.9e-181;
Matches 372; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy          1 MRPGTALQAVLLAVLLVGLRAATGRLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAK 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MRPGTALQAVLLAVLLVGLRAATGRLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAK 60

Qy         61 EACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWT 120
              |||||||||||||||||||| |||||||||||||||||||||||| |||||||
Db         61 EACRRDGGQLVSIESEDEQKLIEXFIENLLPSDGDFWIGLRRREEKQSNSTXCQDLYAWT 120

Qy        121 DGSISQFRNWYVDEPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSD 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 DGSISQFRNWYVDEPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSD 180

Qy        181 EKPAVPSREAEGEETELTTPVLPEETQEEDAKKTFKESREAALNLAYILIPSIPLLLLLV 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 EKPAVPSREAEGEETELTTPVLPEETQEEDAKKTFKESREAALNLAYILIPSIPLLLLLV 240

Qy        241 VTTVVCWVWICRKRKREQPDPSTKKQHTIWPSPHQGNSPDLEVYNVIRKQSEADLAETRP 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 VTTVVCWVWICRKRKREQPDPSTKKQHTIWPSPHQGNSPDLEVYNVIRKQSEADLAETRP 300

Qy        301 DLKNISFRVCSGEATPDDMSCDYDNMAVNPSESGFVTLVSVESGFVTNDIYEFSPDQMGR 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 DLKNISFRVCSGEATPDDMSCDYDNMAVNPSESGFVTLVSVESGFVTNDIYEFSPDQMGR 360

Qy        361 SKESGWVENEIYGY 374
              ||||||||||||||
Db        361 SKESGWVENEIYGY 374
```

RESULT 7
AAY13367
ID AAY13367 standard; protein; 382 AA.
XX
AC AAY13367;
XX
DT 25-JUN-1999 (first entry)
XX
DE Amino acid sequence of protein PRO234.
XX
KW Secreted protein; transmembrane protein; human; enterocolitis;
KW Zollinger-Ellison syndrome; gastrointestinal ulceration;
KW congenital microvillus atrophy; skin disease; cell growth;
KW abnormal keratinocyte differentiation; psoriasis; epithelial cancer;
KW Parkinson's disease; Alzheimer's disease; ALS; neuropathy; fibromodulin;
KW dermal scarring; Usher Syndrome; Atrophia areata; anti-thrombotic;
KW wound healing; tissue repair.
XX
OS Homo sapiens.
XX
PN WO9914328-A2.
XX
PD 25-MAR-1999.
XX
PF 16-SEP-1998; 98WO-US019330.
XX
PR 17-SEP-1997; 97US-0059113P.
PR 17-SEP-1997; 97US-0059115P.
PR 17-SEP-1997; 97US-0059117P.
PR 17-SEP-1997; 97US-0059119P.
PR 17-SEP-1997; 97US-0059121P.
PR 17-SEP-1997; 97US-0059122P.
PR 17-SEP-1997; 97US-0059184P.
PR 18-SEP-1997; 97US-0059263P.
PR 18-SEP-1997; 97US-0059266P.
PR 15-OCT-1997; 97US-0062125P.
PR 17-OCT-1997; 97US-0062285P.
PR 17-OCT-1997; 97US-0062287P.
PR 21-OCT-1997; 97US-0063486P.
PR 24-OCT-1997; 97US-0062814P.
PR 24-OCT-1997; 97US-0062816P.
PR 24-OCT-1997; 97US-0063045P.
PR 24-OCT-1997; 97US-0063120P.
PR 24-OCT-1997; 97US-0063121P.
PR 24-OCT-1997; 97US-0063127P.
PR 24-OCT-1997; 97US-0063128P.
PR 27-OCT-1997; 97US-0063327P.
PR 27-OCT-1997; 97US-0063329P.
PR 28-OCT-1997; 97US-0063541P.
PR 28-OCT-1997; 97US-0063542P.

PR 28-OCT-1997; 97US-0063544P.
PR 28-OCT-1997; 97US-0063549P.
PR 28-OCT-1997; 97US-0063550P.
PR 28-OCT-1997; 97US-0063564P.
PR 29-OCT-1997; 97US-0063435P.
PR 29-OCT-1997; 97US-0063704P.
PR 29-OCT-1997; 97US-0063732P.
PR 29-OCT-1997; 97US-0063734P.
PR 29-OCT-1997; 97US-0063735P.
PR 29-OCT-1997; 97US-0063738P.
PR 29-OCT-1997; 97US-0064215P.
PR 31-OCT-1997; 97US-0063870P.
PR 31-OCT-1997; 97US-0064103P.
PR 03-NOV-1997; 97US-0064248P.
PR 07-NOV-1997; 97US-0064809P.
PR 12-NOV-1997; 97US-0065186P.
PR 17-NOV-1997; 97US-0065846P.
PR 18-NOV-1997; 97US-0065693P.
PR 21-NOV-1997; 97US-0066120P.
PR 21-NOV-1997; 97US-0066364P.
PR 24-NOV-1997; 97US-0066453P.
PR 24-NOV-1997; 97US-0066466P.
PR 24-NOV-1997; 97US-0066511P.
PR 24-NOV-1997; 97US-0066770P.
PR 24-NOV-1997; 97US-0066772P.
PR 25-NOV-1997; 97US-0066840P.
XX
PA (GETH) GENENTECH INC.
XX
PI Wood WI, Gurney AL, Goddard A, Pennica D, Chen J, Yuan J;
XX
DR WPI; 1999-229533/19.
DR N-PSDB; AAX52238.
XX
PT New isolated human genes and polypeptides used in, e.g. treatment of
PT gastrointestinal ulceration.
XX
PS Claim 12; Fig 50; 320pp; English.
XX
CC AAY13344-403 represent secreted and transmembrane human proteins. The
CC cDNA sequences are obtained from cDNA libraries, prepared from fetal
CC lung, fetal kidney, fetal brain, fetal liver and fetal retina. The
CC encoded polypeptides have specific uses based on their homology to known
CC polypeptides, e.g. PRO211 and PRO217 can be used for disorders associated
CC with the preservation and maintenance of gastrointestinal mucosa and the
CC repair of acute and chronic mucosal lesions (e.g. enterocolitis,
CC Zollinger-Ellison syndrome, gastrointestinal ulceration and congenital
CC microvillus atrophy), skin diseases associated with abnormal keratinocyte
CC differentiation (e.g. psoriasis, epithelial cancers such as lung squamous
CC cell carcinoma of the vulva and gliomas), potent effects on cell growth
CC and development, diseases related to growth or survival of nerve cells
CC including Parkinson's disease, Alzheimer's disease, ALS, neuropathies or
CC cancer. PRO265 can be used as for fibromodulin, e.g. for reducing dermal
CC scarring. PRO264 can be used as a target for anti-tumor drugs. PRO533 may
CC be used in the treatment of Usher Syndrome or Atrophia areata; PRO269 can
CC be used as an anti-thrombotic agent; PRO287 polypeptides and portions may
CC have therapeutic applications in wound healing and tissue repair; PRO317
CC can be used for treating problems of the kidney, uterus, endometrium,
CC blood vessels, or related tissue, e.g. in the heart of genital tract
XX
SQ Sequence 382 AA;

Query Match 99.3%; Score 1986; DB 2; Length 382;
Best Local Similarity 97.9%; Pred. No. 4.8e-181;
Matches 374; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

```
Qy          1 MRPGTALQAVLLAVLLVGLRAATGRLLS--------GQPVCRGGTQRPCYKVIYFHDTSR 52
              |||||||||||||||||||||||||||||        ||||||||||||||||||||||||
Db          1 MRPGTALQAVLLAVLLVGLRAATGRLLSASDLDLRGGQPVCRGGTQRPCYKVIYFHDTSR 60

Qy         53 RLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTA 112
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 RLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTA 120

Qy        113 CQDLYAWTDGSISQFRNWYVDEPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKN 172
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CQDLYAWTDGSISQFRNWYVDEPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKN 180

Qy        173 NFICKYSDEKPAVPSREAEGEETELTTPVLPEETQEEDAKKTFKESREAALNLAYILIPS 232
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 NFICKYSDEKPAVPSREAEGEETELTTPVLPEETQEEDAKKTFKESREAALNLAYILIPS 240

Qy        233 IPLLLLLVVTTVVCWVWICRKRKREQPDPSTKKQHTIWPSPHQGNSPDLEVYNVIRKQSE 292
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 IPLLLLLVVTTVVCWVWICRKRKREQPDPSTKKQHTIWPSPHQGNSPDLEVYNVIRKQSE 300

Qy        293 ADLAETRPDLKNISFRVCSGEATPDDMSCDYDNMAVNPSESGFVTLVSVESGFVTNDIYE 352
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ADLAETRPDLKNISFRVCSGEATPDDMSCDYDNMAVNPSESGFVTLVSVESGFVTNDIYE 360

Qy        353 FSPDQMGRSKESGWVENEIYGY 374
              ||||||||||||||||||||||
Db        361 FSPDQMGRSKESGWVENEIYGY 382
```

RESULT 8
ADC78457
ID ADC78457 standard; protein; 382 AA.
XX
AC ADC78457;
XX
DT 01-JAN-2004 (first entry)
XX
DE Human PRO234 protein.
XX
KW antiinflamatory; antiulcer; cytostatic; antipsoriatic; antiparkinsonian;
KW nootropic; neuroprotective; vasotropic; chemotaxic; angiogenic;
KW neurotrophic; osteopathic; antiasthmatic; antiarthritic; antirheumatic;
KW antiarteriosclerotic; cardiant; antidiabetic; cerebroprotective;
KW thrombolytic; immunomodulator; enterocolitis; Zollinger-Ellison syndrome;
KW gastrointestinal ulceration; psoriasis; cancer; Parkinson's disease;
KW Alzheimer's; ALS; neuropathy; dermal scarring; wound healing;
KW nerve repair; thrombosis; bone; cartilage formation; angiogenesis;
KW asthma; rheumatoid arthritis; multiple sclerosis; inflammatory disorder;
KW atherosclerosis; cardiac injury; infertility; premature aging; AIDS;
KW diabetes; stroke; gene therapy; transgenic; PRO; human.
XX
OS Homo sapiens.
XX
PN WO200015796-A2.
XX
PD 23-MAR-2000.
XX
PF 15-SEP-1999; 99WO-US021090.
XX
PR 16-SEP-1998; 98WO-US019330.
XX
PA (GETH) GENENTECH INC.
XX
PI Chen J, Goddard A, Gurney AL, Hillan K, Pennica D, Wood WI;
PI Yuan J;
XX
DR WPI; 2000-271434/23.
DR N-PSDB; ADC78456.
XX
PT Novel nucleic acids encoding secreted and transmembrane polypeptides with
PT homology, e.g. to growth and cancer-associated antigens.
XX
PS Claim 12; SEQ ID NO 137; 355pp; English.
XX
CC The invention relates to a novel nucleic acid encoding a PRO polypeptide.
CC The polypeptides and polynucleotides of the invention may be useful as
CC research tools and as therapeutics for treating enterocolitis, Zollinger-
CC Ellison syndrome, gastrointestinal ulceration, psoriasis, cancer,
CC Parkinson's disease, ALS, neuropathies, dermal
CC scarring and wound healing, nerve repair, thrombosis, bone and/or

```
CC   cartilage formation, angiogenesis, asthma, rheumatoid arthritis, multiple
CC   sclerosis, inflammatory disorders, atherosclerosis, cardiac injury,
CC   infertility, premature aging, AIDS, diabetes complications and stroke.
CC   The molecules may also be utilised during gene therapy procedures and
CC   transgenic animal production. The current sequence is that of the human
CC   PRO protein of the invention.
XX
SQ   Sequence 382 AA;

  Query Match             99.3%;  Score 1986;  DB 3;  Length 382;
  Best Local Similarity   97.9%;  Pred. No. 4.8e-181;
  Matches  374;  Conservative    0;  Mismatches    0;  Indels    8;  Gaps    1;

Qy          1 MRPGTALQAVLLAVLLVGLRAATGRLLS--------GQPVCRGGTQRPCYKVIYFHDTSR 52
              |||||||||||||||||||||||||||        ||||||||||||||||||||||||
Db          1 MRPGTALQAVLLAVLLVGLRAATGRLLSASDLDLRGGQPVCRGGTQRPCYKVIYFHDTSR 60

Qy         53 RLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTA 112
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 RLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTA 120

Qy        113 CQDLYAWTDGSISQFRNWYVDEPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKN 172
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CQDLYAWTDGSISQFRNWYVDEPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKN 180

Qy        173 NFICKYSDEKPAVPSREAEGEETELTTPVLPEETQEEDAKKTFKESREAALNLAYILIPS 232
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 NFICKYSDEKPAVPSREAEGEETELTTPVLPEETQEEDAKKTFKESREAALNLAYILIPS 240

Qy        233 IPLLLLLVVTTVVCWVWICRKRKREQPDPSTKKQHTIWPSPHQGNSPDLEVYNVIRKQSE 292
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 IPLLLLLVVTTVVCWVWICRKRKREQPDPSTKKQHTIWPSPHQGNSPDLEVYNVIRKQSE 300

Qy        293 ADLAETRPDLKNISFRVCSGEATPDDMSCDYDNMAVNPSESGFVTLVSVESGFVTNDIYE 352
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ADLAETRPDLKNISFRVCSGEATPDDMSCDYDNMAVNPSESGFVTLVSVESGFVTNDIYE 360

Qy        353 FSPDQMGRSKESGWVENEIYGY 374
              ||||||||||||||||||||||
Db        361 FSPDQMGRSKESGWVENEIYGY 382


RESULT 9
AAB80235
ID   AAB80235 standard; protein; 382 AA.
XX
AC   AAB80235;
XX
DT   24-APR-2001  (first entry)
XX
DE   Human PRO234 protein.
XX
KW   Human; PRO; dermatological; antipsoriatic; cytostatic; antiinflammatory;
KW   antiparkinsonian nootropic; neuroprotective; vulnerary; cardiant;
KW   antiangiogenic; vasotropic; antiasthmatic; antirheumatic; cancer;
KW   antiarthritic; antiinfertility; antidiabetic; antiviral; diabetes;
KW   ophthalmological; gene therapy; skin disease; gastrointestinal disorder;
KW   ischaemia; inflammation.
XX
OS   Homo sapiens.
XX
PN   WO200104311-A1.
XX
PD   18-JAN-2001.
XX
PF   22-FEB-2000; 2000WO-US004414.
XX
PR   07-JUL-1999;   99US-0143048P.
PR   26-JUL-1999;   99US-0145698P.
PR   28-JUL-1999;   99US-0146222P.
PR   08-SEP-1999;   99WO-US020594.
PR   13-SEP-1999;   99WO-US020944.
PR   15-SEP-1999;   99WO-US021090.
PR   15-SEP-1999;   99WO-US021547.
PR   05-OCT-1999;   99WO-US023089.
PR   29-NOV-1999;   99WO-US028214.
PR   30-NOV-1999;   99WO-US028313.
PR   02-DEC-1999;   99WO-US028564.
PR   02-DEC-1999;   99WO-US028565.
PR   16-DEC-1999;   99WO-US030095.
PR   20-DEC-1999;   99WO-US030911.
PR   20-DEC-1999;   99WO-US030999.
PR   05-JAN-2000; 2000WO-US000219.
XX
PA   (GETH ) GENENTECH INC.
XX
PI   Ashkenazi AJ,  Botstein D,  Desnoyers L,  Eaton DL,  Ferrara N;
PI   Filvaroff E,  Fong S,  Gao W,  Gerber H,  Gerritsen ME,  Goddard A;
PI   Godowski PJ,  Grimaldi CJ,  Gurney AL,  Hillan KJ,  Kljavin IJ;
PI   Mather JP,  Pan J,  Paoni NF,  Roy MA,  Stewart TA,  Tumas D;
PI   Williams PM,  Wood WI;
XX
DR   WPI; 2001-081051/09.
DR   N-PSDB; AAF72396.
XX
PT   Sixty one nucleic acids encoding PRO polypeptides which are useful in the
PT   treatment of skin diseases (e.g. psoriasis), cancers (e.g. lung squamous
PT   cell carcinoma) and neurodegenerative diseases (e.g. Alzheimer's
PT   disease).
XX
PS   Claim 1; Fig 50; 393pp; English.
XX
CC   The present sequence is one of sixty one novel secreted and transmembrane
CC   PRO polypeptides. The PRO polypeptides are useful for treating skin
CC   diseases (e.g. psoriasis), cancers (e.g. lung squamous cell carcinoma),
CC   gastrointestinal disorders (e.g. enterocolitis), neurodegenerative
CC   diseases (e.g. Alzheimer's disease, Parkinson's disease), wound repair,
CC   cardiovascular disorders (e.g. endometrial bleeding angiogenesis,
CC   ischaemias such as coronary ischaemia, atherosclerosis), inflammatory
CC   disorders (e.g. asthma, rheumatoid arthritis, multiple sclerosis),
CC   infertility, AIDS and diabetes and retinal disorders such as retinitis
CC   pigmentosum. The PRO nucleic acids have applications in molecular
CC   biology, including use as hybridization probes, and in chromosome and
CC   gene mapping
XX
SQ   Sequence 382 AA;

  Query Match             99.3%;  Score 1986;  DB 4;  Length 382;
  Best Local Similarity   97.9%;  Pred. No. 4.8e-181;
  Matches  374;  Conservative    0;  Mismatches    0;  Indels    8;  Gaps    1;

Qy          1 MRPGTALQAVLLAVLLVGLRAATGRLLS--------GQPVCRGGTQRPCYKVIYFHDTSR 52
              |||||||||||||||||||||||||||        ||||||||||||||||||||||||
Db          1 MRPGTALQAVLLAVLLVGLRAATGRLLSASDLDLRGGQPVCRGGTQRPCYKVIYFHDTSR 60

Qy         53 RLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTA 112
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 RLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTA 120

Qy        113 CQDLYAWTDGSISQFRNWYVDEPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKN 172
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CQDLYAWTDGSISQFRNWYVDEPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKN 180

Qy        173 NFICKYSDEKPAVPSREAEGEETELTTPVLPEETQEEDAKKTFKESREAALNLAYILIPS 232
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 NFICKYSDEKPAVPSREAEGEETELTTPVLPEETQEEDAKKTFKESREAALNLAYILIPS 240

Qy        233 IPLLLLLVVTTVVCWVWICRKRKREQPDPSTKKQHTIWPSPHQGNSPDLEVYNVIRKQSE 292
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 IPLLLLLVVTTVVCWVWICRKRKREQPDPSTKKQHTIWPSPHQGNSPDLEVYNVIRKQSE 300

Qy        293 ADLAETRPDLKNISFRVCSGEATPDDMSCDYDNMAVNPSESGFVTLVSVESGFVTNDIYE 352
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ADLAETRPDLKNISFRVCSGEATPDDMSCDYDNMAVNPSESGFVTLVSVESGFVTNDIYE 360

Qy        353 FSPDQMGRSKESGWVENEIYGY 374
```

```
                    ||||||||||||||||||||||
Db        361 FSPDQMGRSKESGWVENEIYGY 382


RESULT 10
AAU29033
ID   AAU29033 standard; protein; 382 AA.
XX
AC   AAU29033;
XX
DT   18-DEC-2001  (first entry)
XX
DE   Human PRO polypeptide sequence #10.
XX
KW   PRO polypeptide; mammal; tumour; cancer; human; cattle; horse; sheep;
KW   dog; cat; pig; goat; rabbit; tumour necrosis factor alpha; TNF-alpha;
KW   blood; chondrocyte cell; cell proliferation; cell differentiation; colon;
KW   adrenal; lung; breast; prostate; rectum; cervix; liver; genetic disorder.
XX
OS   Homo sapiens.
XX
PN   WO200168848-A2.
XX
PD   20-SEP-2001.
XX
PF   28-FEB-2001; 2001WO-US006520.
XX
PR   01-MAR-2000; 2000WO-US005601.
PR   02-MAR-2000; 2000WO-US005841.
PR   03-MAR-2000; 2000US-0187202P.
PR   06-MAR-2000; 2000US-0186968P.
PR   14-MAR-2000; 2000US-0189320P.
PR   14-MAR-2000; 2000US-0189328P.
PR   15-MAR-2000; 2000WO-US006884.
PR   21-MAR-2000; 2000US-0190828P.
PR   21-MAR-2000; 2000US-0191007P.
PR   21-MAR-2000; 2000US-0191048P.
PR   21-MAR-2000; 2000US-0191314P.
PR   28-MAR-2000; 2000US-0192655P.
PR   29-MAR-2000; 2000US-0193032P.
PR   29-MAR-2000; 2000US-0193053P.
PR   30-MAR-2000; 2000WO-US008439.
PR   04-APR-2000; 2000US-0194449P.
PR   04-APR-2000; 2000US-0194647P.
PR   11-APR-2000; 2000US-0195975P.
PR   11-APR-2000; 2000US-0196000P.
PR   11-APR-2000; 2000US-0196187P.
PR   11-APR-2000; 2000US-0196690P.
PR   11-APR-2000; 2000US-0196820P.
PR   18-APR-2000; 2000US-0198121P.
PR   18-APR-2000; 2000US-0198585P.
PR   25-APR-2000; 2000US-0199397P.
PR   25-APR-2000; 2000US-0199550P.
PR   25-APR-2000; 2000US-0199654P.
PR   03-MAY-2000; 2000US-0201516P.
PR   17-MAY-2000; 2000WO-US013705.
PR   22-MAY-2000; 2000WO-US014042.
PR   30-MAY-2000; 2000WO-US014941.
PR   02-JUN-2000; 2000WO-US015264.
PR   05-JUN-2000; 2000US-0209832P.
PR   28-JUL-2000; 2000WO-US020710.
PR   22-AUG-2000; 2000US-00644848.
PR   24-AUG-2000; 2000WO-US023328.
PR   08-NOV-2000; 2000WO-US030952.
PR   01-DEC-2000; 2000WO-US032678.
PR   20-DEC-2000; 2000WO-US034956.
XX
PA   (GETH ) GENENTECH INC.
XX
PI   Baker KP,  Chen J,  Desnoyers L,  Goddard A,  Godowski PJ,  Gurney AL;
PI   Pan J,  Smith V,  Watanabe CK,  Wood WI,  Zhang Z;
XX
DR   WPI; 2001-602746/68.
DR   N-PSDB; AAS45934.
XX
PT   Novel nucleic acids encoding PRO polypeptides, used to diagnose the
PT   presence of tumors, such as prostate and breast tumors, in mammals and to
PT   screen for modulators of the compounds.
XX
PS   Claim 11; Fig 20; 774pp; English.
XX
CC   Sequences AAU29024-AAU29328 represent PRO polypeptides of the invention.
CC   The PRO polypeptides and their associated nucleic acids can be used to
CC   detect the presence of a tumour in a mammal by comparing the level of
CC   expression of a PRO polypeptide in a test sample of cells from the animal
CC   and a control sample of normal cells, whereby a higher level of
CC   expression in the test sample indicates the presence of a tumour in the
CC   mammal. Mammals include dogs, cats, cattle, horses, sheep, pigs, goats
CC   and rabbits but are preferably human. The polypeptides can be used to
CC   stimulate tumour necrosis factor (TNF) alpha release from human blood,
CC   when contacted with it. A specific polypeptide can be used to stimulate
CC   the proliferation or differentiation of chondrocyte cells. The PRO
CC   proteins can be used to determine the presence of tumours and also
CC   susceptibility to tumour development, particularly adrenal, lung, colon,
CC   breast, prostate, rectal, cervical, or liver tumours, in mammalian
CC   subjects. The oligonucleotide probes specific for the PRO nucleic acids
CC   can be used for genetic analysis of individuals with genetic disorders
XX
SQ   Sequence 382 AA;

  Query Match             99.3%;  Score 1986;  DB 4;  Length 382;
  Best Local Similarity   97.9%;  Pred. No. 4.8e-181;
  Matches  374;  Conservative    0;  Mismatches    0;  Indels    8;  Gaps    1;

Qy          1 MRPGTALQAVLLAVLLVGLRAATGRLLS--------GQPVCRGGTQRPCYKVIYFHDTSR 52
              |||||||||||||||||||||||||||||        ||||||||||||||||||||||||
Db          1 MRPGTALQAVLLAVLLVGLRAATGRLLSASDLDLRGGQPVCRGGTQRPCYKVIYFHDTSR 60

Qy         53 RLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTA 112
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 RLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTA 120

Qy        113 CQDLYAWTDGSISQFRNWYVDEPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKN 172
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CQDLYAWTDGSISQFRNWYVDEPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKN 180

Qy        173 NFICKYSDEKPAVPSREAEGEETELTTPVLPEETQEEDAKKTFKESREAALNLAYILIPS 232
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 NFICKYSDEKPAVPSREAEGEETELTTPVLPEETQEEDAKKTFKESREAALNLAYILIPS 240

Qy        233 IPLLLLLVVTTVVCWVWICRKRKREQPDPSTKKQHTIWPSPHQGNSPDLEVYNVIRKQSE 292
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 IPLLLLLVVTTVVCWVWICRKRKREQPDPSTKKQHTIWPSPHQGNSPDLEVYNVIRKQSE 300

Qy        293 ADLAETRPDLKNISFRVCSGEATPDDMSCDYDNMAVNPSESGFVTLVSVESGFVTNDIYE 352
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ADLAETRPDLKNISFRVCSGEATPDDMSCDYDNMAVNPSESGFVTLVSVESGFVTNDIYE 360

Qy        353 FSPDQMGRSKESGWVENEIYGY 374
              ||||||||||||||||||||||
Db        361 FSPDQMGRSKESGWVENEIYGY 382


RESULT 11
ABU58409
ID   ABU58409 standard; protein; 382 AA.
XX
AC   ABU58409;
XX
DT   15-APR-2003  (first entry)
XX
DE   Human PRO polypeptide #10.
XX
KW   Human; PRO; cytostatic; tumour; cancer; breast; lung; stomach; liver;
KW   dog; cat; cow; horse; sheep; pig; goat; rabbit; ADEPT;
KW   antibody-dependent enzyme mediated prodrug therapy.
```

XX
OS Homo sapiens.
XX
PN US2003027272-A1.
XX
PD 06-FEB-2003.
XX
PF 21-JUN-2002; 2002US-00176492.
XX
PR 18-SEP-1997; 97US-0059263P.
PR 18-SEP-1997; 97US-0059266P.
PR 17-OCT-1997; 97US-0062250P.
PR 21-OCT-1997; 97US-0063486P.
PR 24-OCT-1997; 97US-0063120P.
PR 24-OCT-1997; 97US-0063121P.
PR 28-OCT-1997; 97US-0063540P.
PR 28-OCT-1997; 97US-0063541P.
PR 28-OCT-1997; 97US-0063544P.
PR 28-OCT-1997; 97US-0063564P.
PR 29-OCT-1997; 97US-0063734P.
PR 31-OCT-1997; 97US-0063870P.
PR 31-OCT-1997; 97US-0064103P.
PR 13-NOV-1997; 97US-0065311P.
PR 21-NOV-1997; 97US-0066120P.
PR 24-NOV-1997; 97US-0066466P.
PR 24-NOV-1997; 97US-0066772P.
PR 11-DEC-1997; 97US-0069335P.
PR 12-DEC-1997; 97US-0069425P.
PR 17-DEC-1997; 97US-0069870P.
PR 18-DEC-1997; 97US-0068017P.
PR 10-MAR-1998; 98US-0077450P.
PR 11-MAR-1998; 98US-0077632P.
PR 11-MAR-1998; 98US-0077649P.
PR 20-MAR-1998; 98US-0078886P.
PR 20-MAR-1998; 98US-0078939P.
PR 27-MAR-1998; 98US-0079664P.
PR 27-MAR-1998; 98US-0079786P.
PR 31-MAR-1998; 98US-0080107P.
PR 31-MAR-1998; 98US-0080194P.
PR 01-APR-1998; 98US-0080327P.
PR 01-APR-1998; 98US-0080333P.
PR 08-APR-1998; 98US-0081049P.
PR 08-APR-1998; 98US-0081070P.
PR 09-APR-1998; 98US-0081195P.
PR 15-APR-1998; 98US-0081838P.
PR 21-APR-1998; 98US-0082568P.
PR 21-APR-1998; 98US-0082569P.
PR 22-APR-1998; 98US-0082704P.
PR 22-APR-1998; 98US-0082797P.
PR 28-APR-1998; 98US-0083322P.
PR 29-APR-1998; 98US-0083495P.
PR 29-APR-1998; 98US-0083496P.
PR 29-APR-1998; 98US-0083499P.
PR 29-APR-1998; 98US-0083559P.
PR 05-MAY-1998; 98US-0084366P.
PR 06-MAY-1998; 98US-0084414P.
PR 07-MAY-1998; 98US-0084639P.
PR 07-MAY-1998; 98US-0084640P.
PR 07-MAY-1998; 98US-0084643P.
PR 15-MAY-1998; 98US-0085579P.
PR 15-MAY-1998; 98US-0085580P.
PR 15-MAY-1998; 98US-0085582P.
PR 15-MAY-1998; 98US-0085700P.
PR 18-MAY-1998; 98US-0086023P.
PR 22-MAY-1998; 98US-0086392P.
PR 22-MAY-1998; 98US-0086486P.
PR 28-MAY-1998; 98US-0087098P.
PR 28-MAY-1998; 98US-0087208P.
PR 02-JUN-1998; 98US-0087609P.
PR 02-JUN-1998; 98US-0087759P.
PR 03-JUN-1998; 98US-0087827P.
PR 04-JUN-1998; 98US-0088025P.
PR 04-JUN-1998; 98US-0088028P.
PR 04-JUN-1998; 98US-0088029P.
PR 04-JUN-1998; 98US-0088033P.
PR 04-JUN-1998; 98US-0088326P.
PR 05-JUN-1998; 98US-0088167P.
PR 05-JUN-1998; 98US-0088202P.
PR 05-JUN-1998; 98US-0088212P.
PR 05-JUN-1998; 98US-0088217P.
PR 09-JUN-1998; 98US-0088655P.
PR 10-JUN-1998; 98US-0088722P.
PR 10-JUN-1998; 98US-0088738P.
PR 10-JUN-1998; 98US-0088740P.
PR 10-JUN-1998; 98US-0088811P.
PR 10-JUN-1998; 98US-0088824P.
PR 10-JUN-1998; 98US-0088825P.
PR 10-JUN-1998; 98US-0088826P.
PR 11-JUN-1998; 98US-0088861P.
PR 11-JUN-1998; 98US-0088863P.
PR 11-JUN-1998; 98US-0088876P.
PR 12-JUN-1998; 98US-0089090P.
PR 12-JUN-1998; 98US-0089105P.
PR 16-JUN-1998; 98US-0089512P.
PR 16-JUN-1998; 98US-0089514P.
PR 17-JUN-1998; 98US-0089538P.
PR 17-JUN-1998; 98US-0089598P.
PR 17-JUN-1998; 98US-0089653P.
PR 18-JUN-1998; 98US-0089908P.
PR 19-JUN-1998; 98US-0089952P.
PR 22-JUN-1998; 98US-0090246P.
PR 22-JUN-1998; 98US-0090252P.
PR 22-JUN-1998; 98US-0090254P.
PR 24-JUN-1998; 98US-0090429P.
PR 24-JUN-1998; 98US-0090435P.
PR 24-JUN-1998; 98US-0090444P.
PR 24-JUN-1998; 98US-0090461P.
PR 24-JUN-1998; 98US-0090535P.
PR 24-JUN-1998; 98US-0090540P.
PR 25-JUN-1998; 98US-0090676P.
PR 25-JUN-1998; 98US-0090678P.
PR 25-JUN-1998; 98US-0090688P.
PR 25-JUN-1998; 98US-0090690P.
PR 25-JUN-1998; 98US-0090694P.
PR 25-JUN-1998; 98US-0090695P.
PR 25-JUN-1998; 98US-0090696P.
PR 26-JUN-1998; 98US-00105413.
PR 26-JUN-1998; 98US-0090862P.
PR 26-JUN-1998; 98US-0090863P.
PR 26-JUN-1998; 98US-0091010P.
PR 01-JUL-1998; 98US-0091359P.
PR 01-JUL-1998; 98US-0091544P.
PR 02-JUL-1998; 98US-0091478P.
PR 02-JUL-1998; 98US-0091486P.
PR 02-JUL-1998; 98US-0091626P.
PR 02-JUL-1998; 98US-0091628P.
PR 02-JUL-1998; 98US-0091632P.
PR 24-JUL-1998; 98US-0094006P.
PR 04-AUG-1998; 98US-0095282P.
PR 10-AUG-1998; 98US-0095998P.
PR 10-AUG-1998; 98US-0096012P.
PR 17-AUG-1998; 98US-0096757P.
PR 17-AUG-1998; 98US-0096766P.
PR 17-AUG-1998; 98US-0096867P.
PR 17-AUG-1998; 98US-0096891P.
PR 17-AUG-1998; 98US-0096897P.
PR 18-AUG-1998; 98US-0096949P.
PR 18-AUG-1998; 98US-0096959P.
PR 18-AUG-1998; 98US-0097022P.
PR 26-AUG-1998; 98US-0097952P.
PR 26-AUG-1998; 98US-0097954P.
PR 26-AUG-1998; 98US-0097955P.
PR 26-AUG-1998; 98US-0097971P.
PR 26-AUG-1998; 98US-0097974P.
PR 26-AUG-1998; 98US-0098014P.
PR 01-SEP-1998; 98US-0098716P.

PR 01-SEP-1998; 98US-0098723P.
PR 02-SEP-1998; 98US-0098803P.
PR 02-SEP-1998; 98US-0098821P.
PR 02-SEP-1998; 98US-0098843P.
PR 09-SEP-1998; 98US-0099602P.
PR 10-SEP-1998; 98US-0099741P.
PR 10-SEP-1998; 98US-0099754P.
PR 10-SEP-1998; 98US-0099763P.
PR 10-SEP-1998; 98US-0099812P.
PR 15-SEP-1998; 98US-0100388P.
PR 16-SEP-1998; 98US-0100662P.
PR 16-SEP-1998; 98US-0100664P.
PR 16-SEP-1998; 98US-0101751P.
PR 16-SEP-1998; 98WO-US019330.
PR 17-SEP-1998; 98US-0100683P.
PR 17-SEP-1998; 98US-0100684P.
PR 17-SEP-1998; 98US-0100919P.
PR 17-SEP-1998; 98US-0100930P.
PR 18-SEP-1998; 98US-0100849P.
PR 18-SEP-1998; 98US-0101014P.
PR 18-SEP-1998; 98US-0101068P.
PR 23-SEP-1998; 98US-0101471P.
PR 23-SEP-1998; 98US-0101472P.
PR 23-SEP-1998; 98US-0101475P.
PR 23-SEP-1998; 98US-0101477P.
PR 24-SEP-1998; 98US-0101738P.
PR 24-SEP-1998; 98US-0101739P.
PR 24-SEP-1998; 98US-0101743P.
PR 24-SEP-1998; 98US-0101922P.
PR 25-SEP-1998; 98US-0101786P.
PR 29-SEP-1998; 98US-0102207P.
PR 29-SEP-1998; 98US-0102240P.
PR 29-SEP-1998; 98US-0102330P.
PR 29-SEP-1998; 98US-0102331P.
PR 30-SEP-1998; 98US-0102487P.
PR 30-SEP-1998; 98US-0102570P.
PR 30-SEP-1998; 98US-0102571P.
PR 01-OCT-1998; 98US-0102684P.
PR 01-OCT-1998; 98US-0102687P.
PR 02-OCT-1998; 98US-0102965P.
PR 06-OCT-1998; 98US-0103258P.
PR 06-OCT-1998; 98US-0103449P.
PR 07-OCT-1998; 98US-00168978.

Query Match 99.3%; Score 1986; DB 6; Length 382;
Best Local Similarity 97.9%; Pred. No. 4.8e-181;
Matches 374; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

```
Qy          1 MRPGTALQAVLLAVLLVGLRAATGRLLS--------GQPVCRGGTQRPCYKVIYFHDTSR 52
              |||||||||||||||||||||||||||        ||||||||||||||||||||||||
Db          1 MRPGTALQAVLLAVLLVGLRAATGRLLSASDLDLRGGQPVCRGGTQRPCYKVIYFHDTSR 60

Qy         53 RLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTA 112
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 RLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTA 120

Qy        113 CQDLYAWTDGSISQFRNWYVDEPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKN 172
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CQDLYAWTDGSISQFRNWYVDEPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKN 180

Qy        173 NFICKYSDEKPAVPSREAEGEETELTTPVLPEETQEEDAKKTFKESREAALNLAYILIPS 232
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 NFICKYSDEKPAVPSREAEGEETELTTPVLPEETQEEDAKKTFKESREAALNLAYILIPS 240

Qy        233 IPLLLLLVVTTVVCWVWICRKRKREQPDPSTKKQHTIWPSPHQGNSPDLEVYNVIRKQSE 292
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 IPLLLLLVVTTVVCWVWICRKRKREQPDPSTKKQHTIWPSPHQGNSPDLEVYNVIRKQSE 300

Qy        293 ADLAETRPDLKNISFRVCSGEATPDDMSCDYDNMAVNPSESGFVTLVSVESGFVTNDIYE 352
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ADLAETRPDLKNISFRVCSGEATPDDMSCDYDNMAVNPSESGFVTLVSVESGFVTNDIYE 360

Qy        353 FSPDQMGRSKESGWVENEIYGY 374
```

```
              ||||||||||||||||||||||
Db        361 FSPDQMGRSKESGWVENEIYGY 382
```

RESULT 12
ABU71613
ID ABU71613 standard; protein; 382 AA.
XX
AC ABU71613;
XX
DT 16-JUN-2003 (first entry)
XX
DE Human PRO polypeptide #24.
XX
KW Human; PRO; secreted polypeptide; transmembrane polypeptide;
KW pathological disorder; cardiac insufficiency disorder; protein secretion;
KW pancreas; diabetes; gastrointestinal mucosa; mucosal lesion; psoriasis;
KW skin disease; keratinocyte differentiation; epithelial cancer; tumour;
KW lung squamous cell carcinoma; epidermoid carcinoma; vulva; glioma;
KW cytostatic; cardiant; endocrine; antidiabetic; gastrointestinal;
KW antiulcer; dermatological; vulnerary.
XX
OS Homo sapiens.
XX
PN US2002146709-A1.
XX
PD 10-OCT-2002.
XX
PF 18-JUL-2001; 2001US-00909088.
XX
PR 17-SEP-1997; 97US-0059113P.
PR 17-SEP-1997; 97US-0059115P.
PR 17-SEP-1997; 97US-0059117P.
PR 17-SEP-1997; 97US-0059119P.
PR 17-SEP-1997; 97US-0059121P.
PR 17-SEP-1997; 97US-0059122P.
PR 17-SEP-1997; 97US-0059184P.
PR 18-SEP-1997; 97US-0059263P.
PR 18-SEP-1997; 97US-0059266P.
PR 15-OCT-1997; 97US-0062125P.
PR 17-OCT-1997; 97US-0062285P.
PR 17-OCT-1997; 97US-0062287P.
PR 21-OCT-1997; 97US-0063486P.
PR 24-OCT-1997; 97US-0062814P.
PR 24-OCT-1997; 97US-0062816P.
PR 24-OCT-1997; 97US-0063045P.
PR 24-OCT-1997; 97US-0063120P.
PR 24-OCT-1997; 97US-0063121P.
PR 24-OCT-1997; 97US-0063127P.
PR 24-OCT-1997; 97US-0063128P.
PR 27-OCT-1997; 97US-0063327P.
PR 27-OCT-1997; 97US-0063329P.
PR 28-OCT-1997; 97US-0063541P.
PR 28-OCT-1997; 97US-0063542P.
PR 28-OCT-1997; 97US-0063544P.
PR 28-OCT-1997; 97US-0063549P.
PR 28-OCT-1997; 97US-0063550P.
PR 28-OCT-1997; 97US-0063564P.
PR 29-OCT-1997; 97US-0063435P.
PR 29-OCT-1997; 97US-0063704P.
PR 29-OCT-1997; 97US-0063732P.
PR 29-OCT-1997; 97US-0063734P.
PR 29-OCT-1997; 97US-0063735P.
PR 29-OCT-1997; 97US-0063738P.
PR 29-OCT-1997; 97US-0064215P.
PR 31-OCT-1997; 97US-0063870P.
PR 31-OCT-1997; 97US-0064103P.
PR 03-NOV-1997; 97US-0064248P.
PR 07-NOV-1997; 97US-0064809P.
PR 12-NOV-1997; 97US-0065186P.
PR 17-NOV-1997; 97US-0065846P.
PR 18-NOV-1997; 97US-0065693P.
PR 21-NOV-1997; 97US-0066120P.

```
PR   21-NOV-1997;   97US-0066364P.
PR   24-NOV-1997;   97US-0066453P.
PR   24-NOV-1997;   97US-0066466P.
PR   24-NOV-1997;   97US-0066511P.
PR   24-NOV-1997;   97US-0066770P.
PR   24-NOV-1997;   97US-0066772P.
PR   10-SEP-1998;   98WO-US018824.
PR   14-SEP-1998;   98WO-US019177.
PR   16-SEP-1998;   98WO-US019330.
PR   17-SEP-1998;   98WO-US019437.
PR   01-DEC-1998;   98WO-US025108.
PR   08-SEP-1999;   99WO-US020594.
PR   13-SEP-1999;   99WO-US020944.
PR   15-SEP-1999;   99WO-US021090.
PR   15-SEP-1999;   99WO-US021547.
PR   05-OCT-1999;   99WO-US023089.
PR   29-NOV-1999;   99WO-US028214.
PR   30-NOV-1999;   99WO-US028313.
PR   01-DEC-1999;   99WO-US028301.
PR   02-DEC-1999;   99WO-US028564.
PR   02-DEC-1999;   99WO-US028565.
PR   16-DEC-1999;   99WO-US030095.
PR   20-DEC-1999;   99WO-US030911.
PR   20-DEC-1999;   99WO-US030999.
PR   05-JAN-2000; 2000WO-US000219.
PR   11-FEB-2000; 2000WO-US003565.
PR   22-FEB-2000; 2000WO-US004414.
PR   24-FEB-2000; 2000WO-US005004.
PR   02-MAR-2000; 2000WO-US005841.
PR   20-MAR-2000; 2000WO-US007377.
PR   30-MAR-2000; 2000WO-US008439.
PR   22-MAY-2000; 2000WO-US014042.
PR   02-JUN-2000; 2000WO-US015264.
PR   28-JUL-2000; 2000WO-US020710.
PR   24-AUG-2000; 2000WO-US023328.
PR   18-SEP-2000; 2000US-00665350.
XX
PA   (GETH ) GENENTECH INC.
XX
PI   Ashkenazi A,  Botstein D,  Desnoyers L,  Eaton DL,  Ferrara N;
PI   Filvaroff E,  Fong S,  Gao W,  Gerber H,  Gerritsen ME,  Goddard A;
PI   Godowski PJ,  Grimaldi JC,  Gurney AL,  Hillan KJ,  Kljavin IJ;
PI   Mather JP,  Pan J,  Paoni NF,  Roy MA,  Stewart TA,  Tumas D;
PI   Williams PM,  Wood WI;
XX
DR   WPI; 2003-328338/31.
DR   N-PSDB; ACA59008.
XX
PT   Isolated nucleic acid useful for e.g., treating pathological disorders
PT   encodes a secreted or transmembrane protein.
XX
PS   Claim 12; Fig 50; 473pp; English.
XX
CC   The invention relates to human PRO polypeptides (secreted or
CC   transmembrane polypeptides) and the polynucleotides encoding them. The
CC   PRO polypeptides and polynucleotides can be used in treating pathological
CC   disorders and tumours, in therapeutic treatment of cardiac insufficiency
CC   disorders and in therapeutic treatment of disorders involving protein
CC   secretion by the pancreas, including diabetes. They can also be used in
CC   treating disorders associated with the preservation and maintenance of
CC   gastrointestinal mucosa and the repair of acute and chronic mucosal
CC   lesions, and skin diseases associated with abnormal keratinocyte
CC   differentiation (e.g., psoriases, epithelial cancers such as lung
CC   squamous cell carcinoma, epidermoid carcinoma of the vulva and gliomas).
CC   The sequences can be used as molecular markers for protein
CC   electrophoresis purposes and can be utilised in protein-protein binding
CC   assays, biochemical screening assays, immunoassays and cell-based assays.
CC   This sequence represents a human PRO polypeptide of the invention
XX
SQ   Sequence 382 AA;

  Query Match             99.3%;  Score 1986;  DB 6;  Length 382;
  Best Local Similarity   97.9%;  Pred. No. 4.8e-181;
  Matches  374;  Conservative    0;  Mismatches    0;  Indels    8;  Gaps    1;

Qy          1 MRPGTALQAVLLAVLLVGLRAATGRLLS--------GQPVCRGGTQRPCYKVIYFHDTSR 52
              ||||||||||||||||||||||||||||        ||||||||||||||||||||||||
Db          1 MRPGTALQAVLLAVLLVGLRAATGRLLSASDLDLRGGQPVCRGGTQRPCYKVIYFHDTSR 60

Qy         53 RLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTA 112
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 RLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTA 120

Qy        113 CQDLYAWTDGSISQFRNWYVDEPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKN 172
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CQDLYAWTDGSISQFRNWYVDEPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKN 180

Qy        173 NFICKYSDEKPAVPSREAEGEETELTTPVLPEETQEEDAKKTFKESREAALNLAYILIPS 232
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 NFICKYSDEKPAVPSREAEGEETELTTPVLPEETQEEDAKKTFKESREAALNLAYILIPS 240

Qy        233 IPLLLLLVVTTVVCWVWICRKRKREQPDPSTKKQHTIWPSPHQGNSPDLEVYNVIRKQSE 292
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 IPLLLLLVVTTVVCWVWICRKRKREQPDPSTKKQHTIWPSPHQGNSPDLEVYNVIRKQSE 300

Qy        293 ADLAETRPDLKNISFRVCSGEATPDDMSCDYDNMAVNPSESGFVTLVSVESGFVTNDIYE 352
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ADLAETRPDLKNISFRVCSGEATPDDMSCDYDNMAVNPSESGFVTLVSVESGFVTNDIYE 360

Qy        353 FSPDQMGRSKESGWVENEIYGY 374
              ||||||||||||||||||||||
Db        361 FSPDQMGRSKESGWVENEIYGY 382


RESULT 13
ABU87957
ID   ABU87957 standard; protein; 382 AA.
XX
AC   ABU87957;
XX
DT   07-JUL-2003  (first entry)
XX
DE   Novel human secreted and transmembrane protein PRO234.
XX
KW   Human; secreted and transmembrane protein: PRO; gene therapy;
KW   tumour necrosis factor-alpha release; TNF-alpha release;
KW   chondrocyte proliferation; chondrocyte differentiation; tumour;
KW   adrenal tumour; lung tumour; colon tumour; breast tumour;
KW   prostate tumour; rectal tumour; cervical tumour; liver tumour.
XX
OS   Homo sapiens.
XX
PN   US2003032127-A1.
XX
PD   13-FEB-2003.
XX
PF   26-JUN-2002; 2002US-00183012.
XX
PR   18-SEP-1997;   97US-0059263P.
PR   18-SEP-1997;   97US-0059266P.
PR   17-OCT-1997;   97US-0062250P.
PR   21-OCT-1997;   97US-0063486P.
PR   24-OCT-1997;   97US-0063120P.
PR   24-OCT-1997;   97US-0063121P.
PR   28-OCT-1997;   97US-0063540P.
PR   28-OCT-1997;   97US-0063541P.
PR   28-OCT-1997;   97US-0063544P.
PR   28-OCT-1997;   97US-0063564P.
PR   29-OCT-1997;   97US-0063734P.
PR   31-OCT-1997;   97US-0063870P.
PR   31-OCT-1997;   97US-0064103P.
PR   13-NOV-1997;   97US-0065311P.
PR   21-NOV-1997;   97US-0066120P.
PR   24-NOV-1997;   97US-0066466P.
PR   24-NOV-1997;   97US-0066772P.
PR   11-DEC-1997;   97US-0069335P.
```

PR 12-DEC-1997; 97US-0069425P.
PR 17-DEC-1997; 97US-0069870P.
PR 18-DEC-1997; 97US-0068017P.
PR 10-MAR-1998; 98US-0077450P.
PR 11-MAR-1998; 98US-0077632P.
PR 11-MAR-1998; 98US-0077649P.
PR 20-MAR-1998; 98US-0078886P.
PR 20-MAR-1998; 98US-0078939P.
PR 27-MAR-1998; 98US-0079664P.
PR 27-MAR-1998; 98US-0079786P.
PR 31-MAR-1998; 98US-0080107P.
PR 31-MAR-1998; 98US-0080194P.
PR 01-APR-1998; 98US-0080327P.
PR 01-APR-1998; 98US-0080333P.
PR 08-APR-1998; 98US-0081049P.
PR 08-APR-1998; 98US-0081070P.
PR 09-APR-1998; 98US-0081195P.
PR 15-APR-1998; 98US-0081838P.
PR 21-APR-1998; 98US-0082568P.
PR 21-APR-1998; 98US-0082569P.
PR 22-APR-1998; 98US-0082704P.
PR 22-APR-1998; 98US-0082797P.
PR 28-APR-1998; 98US-0083322P.
PR 29-APR-1998; 98US-0083495P.
PR 29-APR-1998; 98US-0083496P.
PR 29-APR-1998; 98US-0083499P.
PR 29-APR-1998; 98US-0083559P.
PR 05-MAY-1998; 98US-0084366P.
PR 06-MAY-1998; 98US-0084414P.
PR 07-MAY-1998; 98US-0084639P.
PR 07-MAY-1998; 98US-0084640P.
PR 07-MAY-1998; 98US-0084643P.
PR 15-MAY-1998; 98US-0085579P.
PR 15-MAY-1998; 98US-0085580P.
PR 15-MAY-1998; 98US-0085582P.
PR 15-MAY-1998; 98US-0085700P.
PR 18-MAY-1998; 98US-0086023P.
PR 22-MAY-1998; 98US-0086392P.
PR 22-MAY-1998; 98US-0086486P.
PR 28-MAY-1998; 98US-0087098P.
PR 28-MAY-1998; 98US-0087208P.
PR 02-JUN-1998; 98US-0087609P.
PR 02-JUN-1998; 98US-0087759P.
PR 03-JUN-1998; 98US-0087827P.
PR 04-JUN-1998; 98US-0088025P.
PR 04-JUN-1998; 98US-0088028P.
PR 04-JUN-1998; 98US-0088029P.
PR 04-JUN-1998; 98US-0088033P.
PR 04-JUN-1998; 98US-0088326P.
PR 05-JUN-1998; 98US-0088167P.
PR 05-JUN-1998; 98US-0088202P.
PR 05-JUN-1998; 98US-0088212P.
PR 05-JUN-1998; 98US-0088217P.
PR 09-JUN-1998; 98US-0088655P.
PR 10-JUN-1998; 98US-0088722P.
PR 10-JUN-1998; 98US-0088738P.
PR 10-JUN-1998; 98US-0088740P.
PR 10-JUN-1998; 98US-0088811P.
PR 10-JUN-1998; 98US-0088824P.
PR 10-JUN-1998; 98US-0088825P.
PR 10-JUN-1998; 98US-0088826P.
PR 11-JUN-1998; 98US-0088861P.
PR 11-JUN-1998; 98US-0088863P.
PR 11-JUN-1998; 98US-0088876P.
PR 12-JUN-1998; 98US-0089090P.
PR 12-JUN-1998; 98US-0089105P.
PR 16-JUN-1998; 98US-0089512P.
PR 16-JUN-1998; 98US-0089514P.
PR 17-JUN-1998; 98US-0089538P.
PR 17-JUN-1998; 98US-0089598P.
PR 17-JUN-1998; 98US-0089653P.
PR 18-JUN-1998; 98US-0089908P.
PR 19-JUN-1998; 98US-0089952P.
PR 22-JUN-1998; 98US-0090246P.
PR 22-JUN-1998; 98US-0090252P.
PR 22-JUN-1998; 98US-0090254P.
PR 24-JUN-1998; 98US-0090429P.
PR 24-JUN-1998; 98US-0090435P.
PR 24-JUN-1998; 98US-0090444P.
PR 24-JUN-1998; 98US-0090461P.
PR 24-JUN-1998; 98US-0090535P.
PR 24-JUN-1998; 98US-0090540P.
PR 25-JUN-1998; 98US-0090676P.
PR 25-JUN-1998; 98US-0090678P.
PR 25-JUN-1998; 98US-0090688P.
PR 25-JUN-1998; 98US-0090690P.
PR 25-JUN-1998; 98US-0090694P.
PR 25-JUN-1998; 98US-0090695P.
PR 25-JUN-1998; 98US-0090696P.
PR 26-JUN-1998; 98US-00105413.
PR 26-JUN-1998; 98US-0090862P.
PR 26-JUN-1998; 98US-0090863P.
PR 26-JUN-1998; 98US-0091010P.
PR 01-JUL-1998; 98US-0091359P.
PR 01-JUL-1998; 98US-0091544P.
PR 02-JUL-1998; 98US-0091478P.
PR 02-JUL-1998; 98US-0091486P.
PR 02-JUL-1998; 98US-0091626P.
PR 02-JUL-1998; 98US-0091628P.
PR 02-JUL-1998; 98US-0091632P.
PR 24-JUL-1998; 98US-0094006P.
PR 04-AUG-1998; 98US-0095282P.
PR 10-AUG-1998; 98US-0095998P.
PR 10-AUG-1998; 98US-0096012P.
PR 17-AUG-1998; 98US-0096757P.
PR 17-AUG-1998; 98US-0096766P.
PR 17-AUG-1998; 98US-0096867P.
PR 17-AUG-1998; 98US-0096891P.
PR 17-AUG-1998; 98US-0096897P.
PR 18-AUG-1998; 98US-0096949P.
PR 18-AUG-1998; 98US-0096959P.
PR 18-AUG-1998; 98US-0097022P.
PR 26-AUG-1998; 98US-0097952P.
PR 26-AUG-1998; 98US-0097954P.
PR 26-AUG-1998; 98US-0097955P.
PR 26-AUG-1998; 98US-0097971P.
PR 26-AUG-1998; 98US-0097974P.
PR 26-AUG-1998; 98US-0098014P.
PR 01-SEP-1998; 98US-0098716P.
PR 01-SEP-1998; 98US-0098723P.
PR 02-SEP-1998; 98US-0098803P.
PR 02-SEP-1998; 98US-0098821P.
PR 02-SEP-1998; 98US-0098843P.
PR 09-SEP-1998; 98US-0099602P.
PR 10-SEP-1998; 98US-0099741P.
PR 10-SEP-1998; 98US-0099754P.
PR 10-SEP-1998; 98US-0099763P.
PR 10-SEP-1998; 98US-0099812P.
PR 15-SEP-1998; 98US-0100388P.
PR 16-SEP-1998; 98US-0100662P.
PR 16-SEP-1998; 98US-0100664P.
PR 16-SEP-1998; 98US-0101751P.
PR 16-SEP-1998; 98WO-US019330.
PR 17-SEP-1998; 98US-0100683P.
PR 17-SEP-1998; 98US-0100684P.
PR 17-SEP-1998; 98US-0100919P.
PR 17-SEP-1998; 98US-0100930P.
PR 18-SEP-1998; 98US-0100849P.
PR 18-SEP-1998; 98US-0101014P.
PR 18-SEP-1998; 98US-0101068P.
PR 23-SEP-1998; 98US-0101471P.
PR 23-SEP-1998; 98US-0101472P.
PR 23-SEP-1998; 98US-0101475P.
PR 23-SEP-1998; 98US-0101477P.
PR 24-SEP-1998; 98US-0101738P.
PR 24-SEP-1998; 98US-0101739P.

PR 24-SEP-1998; 98US-0101743P.
PR 24-SEP-1998; 98US-0101922P.
PR 25-SEP-1998; 98US-0101786P.
PR 29-SEP-1998; 98US-0102207P.
PR 29-SEP-1998; 98US-0102240P.
PR 29-SEP-1998; 98US-0102330P.
PR 29-SEP-1998; 98US-0102331P.
PR 30-SEP-1998; 98US-0102487P.
PR 30-SEP-1998; 98US-0102570P.
PR 30-SEP-1998; 98US-0102571P.
PR 01-OCT-1998; 98US-0102684P.
PR 01-OCT-1998; 98US-0102687P.
PR 02-OCT-1998; 98US-0102965P.
PR 06-OCT-1998; 98US-0103258P.

Query Match 99.3%; Score 1986; DB 6; Length 382;
Best Local Similarity 97.9%; Pred. No. 4.8e-181;
Matches 374; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

```
Qy          1 MRPGTALQAVLLAVLLVGLRAATGRLLS--------GQPVCRGGTQRPCYKVIYFHDTSR 52
              |||||||||||||||||||||||||||        ||||||||||||||||||||||||
Db          1 MRPGTALQAVLLAVLLVGLRAATGRLLSASDLDLRGGQPVCRGGTQRPCYKVIYFHDTSR 60

Qy         53 RLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTA 112
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 RLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTA 120

Qy        113 CQDLYAWTDGSISQFRNWYVDEPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKN 172
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CQDLYAWTDGSISQFRNWYVDEPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKN 180

Qy        173 NFICKYSDEKPAVPSREAEGEETELTTPVLPEETQEEDAKKTFKESREAALNLAYILIPS 232
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 NFICKYSDEKPAVPSREAEGEETELTTPVLPEETQEEDAKKTFKESREAALNLAYILIPS 240

Qy        233 IPLLLLLVVTTVVCWVWICRKRKREQPDPSTKKQHTIWPSPHQGNSPDLEVYNVIRKQSE 292
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 IPLLLLLVVTTVVCWVWICRKRKREQPDPSTKKQHTIWPSPHQGNSPDLEVYNVIRKQSE 300

Qy        293 ADLAETRPDLKNISFRVCSGEATPDDMSCDYDNMAVNPSESGFVTLVSVESGFVTNDIYE 352
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ADLAETRPDLKNISFRVCSGEATPDDMSCDYDNMAVNPSESGFVTLVSVESGFVTNDIYE 360

Qy        353 FSPDQMGRSKESGWVENEIYGY 374
              ||||||||||||||||||||||
Db        361 FSPDQMGRSKESGWVENEIYGY 382
```

RESULT 14
ABU84272
ID ABU84272 standard; protein; 382 AA.
XX
AC ABU84272;
XX
DT 02-AUG-2003 (first entry)
XX
DE Human secreted/transmembrane protein (PRO) #10.
XX
KW Human; secreted and transmembrane protein; PRO; TNF-alpha;
KW tumour necrosis factor alpha; chondrocyte cell; tumour; gene therapy;
KW tissue typing.
XX
OS Homo sapiens.
XX
PN US2003032112-A1.
XX
PD 13-FEB-2003.
XX
PF 21-JUN-2002; 2002US-00176756.
XX
PR 18-SEP-1997; 97US-0059263P.
PR 18-SEP-1997; 97US-0059266P.
PR 17-OCT-1997; 97US-0062250P.
PR 21-OCT-1997; 97US-0063486P.
PR 24-OCT-1997; 97US-0063120P.
PR 24-OCT-1997; 97US-0063121P.
PR 28-OCT-1997; 97US-0063540P.
PR 28-OCT-1997; 97US-0063541P.
PR 28-OCT-1997; 97US-0063544P.
PR 28-OCT-1997; 97US-0063564P.
PR 29-OCT-1997; 97US-0063734P.
PR 31-OCT-1997; 97US-0063870P.
PR 31-OCT-1997; 97US-0064103P.
PR 13-NOV-1997; 97US-0065311P.
PR 21-NOV-1997; 97US-0066120P.
PR 24-NOV-1997; 97US-0066466P.
PR 24-NOV-1997; 97US-0066772P.
PR 11-DEC-1997; 97US-0069335P.
PR 12-DEC-1997; 97US-0069425P.
PR 17-DEC-1997; 97US-0069870P.
PR 18-DEC-1997; 97US-0068017P.
PR 10-MAR-1998; 98US-0077450P.
PR 11-MAR-1998; 98US-0077632P.
PR 11-MAR-1998; 98US-0077649P.
PR 20-MAR-1998; 98US-0078886P.
PR 20-MAR-1998; 98US-0078939P.
PR 27-MAR-1998; 98US-0079664P.
PR 27-MAR-1998; 98US-0079786P.
PR 31-MAR-1998; 98US-0080107P.
PR 31-MAR-1998; 98US-0080194P.
PR 01-APR-1998; 98US-0080327P.
PR 01-APR-1998; 98US-0080333P.
PR 08-APR-1998; 98US-0081049P.
PR 08-APR-1998; 98US-0081070P.
PR 09-APR-1998; 98US-0081195P.
PR 15-APR-1998; 98US-0081838P.
PR 21-APR-1998; 98US-0082568P.
PR 21-APR-1998; 98US-0082569P.
PR 22-APR-1998; 98US-0082704P.
PR 22-APR-1998; 98US-0082797P.
PR 28-APR-1998; 98US-0083322P.
PR 29-APR-1998; 98US-0083495P.
PR 29-APR-1998; 98US-0083496P.
PR 29-APR-1998; 98US-0083499P.
PR 29-APR-1998; 98US-0083559P.
PR 05-MAY-1998; 98US-0084366P.
PR 06-MAY-1998; 98US-0084414P.
PR 07-MAY-1998; 98US-0084639P.
PR 07-MAY-1998; 98US-0084640P.
PR 07-MAY-1998; 98US-0084643P.
PR 15-MAY-1998; 98US-0085579P.
PR 15-MAY-1998; 98US-0085580P.
PR 15-MAY-1998; 98US-0085582P.
PR 15-MAY-1998; 98US-0085700P.
PR 18-MAY-1998; 98US-0086023P.
PR 22-MAY-1998; 98US-0086392P.
PR 22-MAY-1998; 98US-0086486P.
PR 28-MAY-1998; 98US-0087098P.
PR 28-MAY-1998; 98US-0087208P.
PR 02-JUN-1998; 98US-0087609P.
PR 02-JUN-1998; 98US-0087759P.
PR 03-JUN-1998; 98US-0087827P.
PR 04-JUN-1998; 98US-0088025P.
PR 04-JUN-1998; 98US-0088028P.
PR 04-JUN-1998; 98US-0088029P.
PR 04-JUN-1998; 98US-0088033P.
PR 04-JUN-1998; 98US-0088326P.
PR 05-JUN-1998; 98US-0088167P.
PR 05-JUN-1998; 98US-0088202P.
PR 05-JUN-1998; 98US-0088212P.
PR 05-JUN-1998; 98US-0088217P.
PR 09-JUN-1998; 98US-0088655P.
PR 10-JUN-1998; 98US-0088722P.
PR 10-JUN-1998; 98US-0088738P.
PR 10-JUN-1998; 98US-0088740P.
PR 10-JUN-1998; 98US-0088811P.

PR 10-JUN-1998; 98US-0088824P.
PR 10-JUN-1998; 98US-0088825P.
PR 10-JUN-1998; 98US-0088826P.
PR 11-JUN-1998; 98US-0088861P.
PR 11-JUN-1998; 98US-0088863P.
PR 11-JUN-1998; 98US-0088876P.
PR 12-JUN-1998; 98US-0089090P.
PR 12-JUN-1998; 98US-0089105P.
PR 16-JUN-1998; 98US-0089512P.
PR 16-JUN-1998; 98US-0089514P.
PR 17-JUN-1998; 98US-0089538P.
PR 17-JUN-1998; 98US-0089598P.
PR 17-JUN-1998; 98US-0089653P.
PR 18-JUN-1998; 98US-0089908P.
PR 19-JUN-1998; 98US-0089952P.
PR 22-JUN-1998; 98US-0090246P.
PR 22-JUN-1998; 98US-0090252P.
PR 22-JUN-1998; 98US-0090254P.
PR 24-JUN-1998; 98US-0090429P.
PR 24-JUN-1998; 98US-0090435P.
PR 24-JUN-1998; 98US-0090444P.
PR 24-JUN-1998; 98US-0090461P.
PR 24-JUN-1998; 98US-0090535P.
PR 24-JUN-1998; 98US-0090540P.
PR 25-JUN-1998; 98US-0090676P.
PR 25-JUN-1998; 98US-0090678P.
PR 25-JUN-1998; 98US-0090688P.
PR 25-JUN-1998; 98US-0090690P.
PR 25-JUN-1998; 98US-0090694P.
PR 25-JUN-1998; 98US-0090695P.
PR 25-JUN-1998; 98US-0090696P.
PR 26-JUN-1998; 98US-00105413.
PR 26-JUN-1998; 98US-0090862P.
PR 26-JUN-1998; 98US-0090863P.
PR 26-JUN-1998; 98US-0091010P.
PR 01-JUL-1998; 98US-0091359P.
PR 01-JUL-1998; 98US-0091544P.
PR 02-JUL-1998; 98US-0091478P.
PR 02-JUL-1998; 98US-0091486P.
PR 02-JUL-1998; 98US-0091626P.
PR 02-JUL-1998; 98US-0091628P.
PR 02-JUL-1998; 98US-0091632P.
PR 24-JUL-1998; 98US-0094006P.
PR 04-AUG-1998; 98US-0095282P.
PR 10-AUG-1998; 98US-0095998P.
PR 10-AUG-1998; 98US-0096012P.
PR 17-AUG-1998; 98US-0096757P.
PR 17-AUG-1998; 98US-0096766P.
PR 17-AUG-1998; 98US-0096867P.
PR 17-AUG-1998; 98US-0096891P.
PR 17-AUG-1998; 98US-0096897P.
PR 18-AUG-1998; 98US-0096949P.
PR 18-AUG-1998; 98US-0096959P.
PR 18-AUG-1998; 98US-0097022P.
PR 26-AUG-1998; 98US-0097952P.
PR 26-AUG-1998; 98US-0097954P.
PR 26-AUG-1998; 98US-0097955P.
PR 26-AUG-1998; 98US-0097971P.
PR 26-AUG-1998; 98US-0097974P.
PR 26-AUG-1998; 98US-0098014P.
PR 01-SEP-1998; 98US-0098716P.
PR 01-SEP-1998; 98US-0098723P.
PR 02-SEP-1998; 98US-0098803P.
PR 02-SEP-1998; 98US-0098821P.
PR 02-SEP-1998; 98US-0098843P.
PR 09-SEP-1998; 98US-0099602P.
PR 10-SEP-1998; 98US-0099741P.
PR 10-SEP-1998; 98US-0099754P.
PR 10-SEP-1998; 98US-0099763P.
PR 10-SEP-1998; 98US-0099812P.
PR 15-SEP-1998; 98US-0100388P.
PR 16-SEP-1998; 98US-0100662P.
PR 16-SEP-1998; 98US-0100664P.
PR 16-SEP-1998; 98US-0101751P.
PR 16-SEP-1998; 98WO-US019330.
PR 17-SEP-1998; 98US-0100683P.
PR 17-SEP-1998; 98US-0100684P.
PR 17-SEP-1998; 98US-0100919P.
PR 17-SEP-1998; 98US-0100930P.
PR 18-SEP-1998; 98US-0100849P.
PR 18-SEP-1998; 98US-0101014P.
PR 18-SEP-1998; 98US-0101068P.
PR 23-SEP-1998; 98US-0101471P.
PR 23-SEP-1998; 98US-0101472P.
PR 23-SEP-1998; 98US-0101475P.
PR 23-SEP-1998; 98US-0101477P.
PR 24-SEP-1998; 98US-0101738P.
PR 24-SEP-1998; 98US-0101739P.
PR 24-SEP-1998; 98US-0101743P.
PR 24-SEP-1998; 98US-0101922P.
PR 25-SEP-1998; 98US-0101786P.
PR 29-SEP-1998; 98US-0102207P.
PR 29-SEP-1998; 98US-0102240P.
PR 29-SEP-1998; 98US-0102330P.
PR 29-SEP-1998; 98US-0102331P.
PR 30-SEP-1998; 98US-0102487P.
PR 30-SEP-1998; 98US-0102570P.
PR 30-SEP-1998; 98US-0102571P.
PR 01-OCT-1998; 98US-0102684P.
PR 01-OCT-1998; 98US-0102687P.
PR 02-OCT-1998; 98US-0102965P.
PR 06-OCT-1998; 98US-0103258P.
PR 06-OCT-1998; 98US-0103449P.
PR 07-OCT-1998; 98US-00168978.

Query Match 99.3%; Score 1986; DB 6; Length 382;
Best Local Similarity 97.9%; Pred. No. 4.8e-181;
Matches 374; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

```
Qy          1 MRPGTALQAVLLAVLLVGLRAATGRLLS--------GQPVCRGGTQRPCYKVIYFHDTSR 52
              |||||||||||||||||||||||||||||        ||||||||||||||||||||||||
Db          1 MRPGTALQAVLLAVLLVGLRAATGRLLSASDLDLRGGQPVCRGGTQRPCYKVIYFHDTSR 60

Qy         53 RLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTA 112
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 RLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTA 120

Qy        113 CQDLYAWTDGSISQFRNWYVDEPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKN 172
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CQDLYAWTDGSISQFRNWYVDEPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKN 180

Qy        173 NFICKYSDEKPAVPSREAEGEETELTTPVLPEETQEEDAKKTFKESREAALNLAYILIPS 232
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 NFICKYSDEKPAVPSREAEGEETELTTPVLPEETQEEDAKKTFKESREAALNLAYILIPS 240

Qy        233 IPLLLLLVVTTVVCWVWICRKRKREQPDPSTKKQHTIWPSPHQGNSPDLEVYNVIRKQSE 292
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 IPLLLLLVVTTVVCWVWICRKRKREQPDPSTKKQHTIWPSPHQGNSPDLEVYNVIRKQSE 300

Qy        293 ADLAETRPDLKNISFRVCSGEATPDDMSCDYDNMAVNPSESGFVTLVSVESGFVTNDIYE 352
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ADLAETRPDLKNISFRVCSGEATPDDMSCDYDNMAVNPSESGFVTLVSVESGFVTNDIYE 360

Qy        353 FSPDQMGRSKESGWVENEIYGY 374
              ||||||||||||||||||||||
Db        361 FSPDQMGRSKESGWVENEIYGY 382
```

RESULT 15
ABR66146
ID ABR66146 standard; protein; 382 AA.
XX
AC ABR66146;
XX
DT 05-AUG-2003 (first entry)
XX

DE Human secreted polypeptide PRO234, SEQ ID NO:20.
XX
KW Human; PRO; secreted protein; transmembrane protein;
KW extracellular domain; tumour necrosis factor-alpha; TNF-alpha;
KW chondrocyte; proliferation; differentiation; cartilage disorder;
KW bone disorder; arthritis; sports injury; cancer; tumour; diagnosis;
KW adrenal tumour; lung; colon; breast; prostate; kidney; rectum; cervix;
KW liver; drug screening; transgenic animal; genetic analysis;
KW antiarthritic; vulnerary; gene therapy.
XX
OS Homo sapiens.
XX
PN US2003027278-A1.
XX
PD 06-FEB-2003.
XX
PF 21-JUN-2002; 2002US-00176987.
XX
PR 18-SEP-1997; 97US-0059263P.
PR 18-SEP-1997; 97US-0059266P.
PR 17-OCT-1997; 97US-0062250P.
PR 21-OCT-1997; 97US-0063486P.
PR 24-OCT-1997; 97US-0063120P.
PR 24-OCT-1997; 97US-0063121P.
PR 28-OCT-1997; 97US-0063540P.
PR 28-OCT-1997; 97US-0063541P.
PR 28-OCT-1997; 97US-0063544P.
PR 28-OCT-1997; 97US-0063564P.
PR 29-OCT-1997; 97US-0063734P.
PR 31-OCT-1997; 97US-0063870P.
PR 31-OCT-1997; 97US-0064103P.
PR 13-NOV-1997; 97US-0065311P.
PR 21-NOV-1997; 97US-0066120P.
PR 24-NOV-1997; 97US-0066466P.
PR 24-NOV-1997; 97US-0066772P.
PR 11-DEC-1997; 97US-0069335P.
PR 12-DEC-1997; 97US-0069425P.
PR 17-DEC-1997; 97US-0069870P.
PR 18-DEC-1997; 97US-0068017P.
PR 10-MAR-1998; 98US-0077450P.
PR 11-MAR-1998; 98US-0077632P.
PR 11-MAR-1998; 98US-0077649P.
PR 20-MAR-1998; 98US-0078886P.
PR 20-MAR-1998; 98US-0078939P.
PR 27-MAR-1998; 98US-0079664P.
PR 27-MAR-1998; 98US-0079786P.
PR 31-MAR-1998; 98US-0080107P.
PR 31-MAR-1998; 98US-0080194P.
PR 01-APR-1998; 98US-0080327P.
PR 01-APR-1998; 98US-0080333P.
PR 08-APR-1998; 98US-0081049P.
PR 08-APR-1998; 98US-0081070P.
PR 09-APR-1998; 98US-0081195P.
PR 15-APR-1998; 98US-0081838P.
PR 21-APR-1998; 98US-0082568P.
PR 21-APR-1998; 98US-0082569P.
PR 22-APR-1998; 98US-0082704P.
PR 22-APR-1998; 98US-0082797P.
PR 28-APR-1998; 98US-0083322P.
PR 29-APR-1998; 98US-0083495P.
PR 29-APR-1998; 98US-0083496P.
PR 29-APR-1998; 98US-0083499P.
PR 29-APR-1998; 98US-0083559P.
PR 05-MAY-1998; 98US-0084366P.
PR 06-MAY-1998; 98US-0084414P.
PR 07-MAY-1998; 98US-0084639P.
PR 07-MAY-1998; 98US-0084640P.
PR 07-MAY-1998; 98US-0084643P.
PR 15-MAY-1998; 98US-0085579P.
PR 15-MAY-1998; 98US-0085580P.
PR 15-MAY-1998; 98US-0085582P.
PR 15-MAY-1998; 98US-0085700P.
PR 18-MAY-1998; 98US-0086023P.
PR 22-MAY-1998; 98US-0086392P.
PR 22-MAY-1998; 98US-0086486P.
PR 28-MAY-1998; 98US-0087098P.
PR 28-MAY-1998; 98US-0087208P.
PR 02-JUN-1998; 98US-0087609P.
PR 02-JUN-1998; 98US-0087759P.
PR 03-JUN-1998; 98US-0087827P.
PR 04-JUN-1998; 98US-0088025P.
PR 04-JUN-1998; 98US-0088028P.
PR 04-JUN-1998; 98US-0088029P.
PR 04-JUN-1998; 98US-0088033P.
PR 04-JUN-1998; 98US-0088326P.
PR 05-JUN-1998; 98US-0088167P.
PR 05-JUN-1998; 98US-0088202P.
PR 05-JUN-1998; 98US-0088212P.
PR 05-JUN-1998; 98US-0088217P.
PR 09-JUN-1998; 98US-0088655P.
PR 10-JUN-1998; 98US-0088722P.
PR 10-JUN-1998; 98US-0088738P.
PR 10-JUN-1998; 98US-0088740P.
PR 10-JUN-1998; 98US-0088811P.
PR 10-JUN-1998; 98US-0088824P.
PR 10-JUN-1998; 98US-0088825P.
PR 10-JUN-1998; 98US-0088826P.
PR 11-JUN-1998; 98US-0088861P.
PR 11-JUN-1998; 98US-0088863P.
PR 11-JUN-1998; 98US-0088876P.
PR 12-JUN-1998; 98US-0089090P.
PR 12-JUN-1998; 98US-0089105P.
PR 16-JUN-1998; 98US-0089512P.
PR 16-JUN-1998; 98US-0089514P.
PR 17-JUN-1998; 98US-0089538P.
PR 17-JUN-1998; 98US-0089598P.
PR 17-JUN-1998; 98US-0089653P.
PR 18-JUN-1998; 98US-0089908P.
PR 19-JUN-1998; 98US-0089952P.
PR 22-JUN-1998; 98US-0090246P.
PR 22-JUN-1998; 98US-0090252P.
PR 22-JUN-1998; 98US-0090254P.
PR 24-JUN-1998; 98US-0090429P.
PR 24-JUN-1998; 98US-0090435P.
PR 24-JUN-1998; 98US-0090444P.
PR 24-JUN-1998; 98US-0090461P.
PR 24-JUN-1998; 98US-0090535P.
PR 24-JUN-1998; 98US-0090540P.
PR 25-JUN-1998; 98US-0090676P.
PR 25-JUN-1998; 98US-0090678P.
PR 25-JUN-1998; 98US-0090688P.
PR 25-JUN-1998; 98US-0090690P.
PR 25-JUN-1998; 98US-0090694P.
PR 25-JUN-1998; 98US-0090695P.
PR 25-JUN-1998; 98US-0090696P.
PR 26-JUN-1998; 98US-00105413.
PR 26-JUN-1998; 98US-0090862P.
PR 26-JUN-1998; 98US-0090863P.
PR 26-JUN-1998; 98US-0091010P.
PR 01-JUL-1998; 98US-0091359P.
PR 01-JUL-1998; 98US-0091544P.
PR 02-JUL-1998; 98US-0091478P.
PR 02-JUL-1998; 98US-0091486P.
PR 02-JUL-1998; 98US-0091626P.
PR 02-JUL-1998; 98US-0091628P.
PR 02-JUL-1998; 98US-0091632P.
PR 24-JUL-1998; 98US-0094006P.
PR 04-AUG-1998; 98US-0095282P.
PR 10-AUG-1998; 98US-0095998P.
PR 10-AUG-1998; 98US-0096012P.
PR 17-AUG-1998; 98US-0096757P.
PR 17-AUG-1998; 98US-0096766P.
PR 17-AUG-1998; 98US-0096867P.
PR 17-AUG-1998; 98US-0096891P.
PR 17-AUG-1998; 98US-0096897P.
PR 18-AUG-1998; 98US-0096949P.

PR 18-AUG-1998; 98US-0096959P.
PR 18-AUG-1998; 98US-0097022P.
PR 26-AUG-1998; 98US-0097952P.
PR 26-AUG-1998; 98US-0097954P.
PR 26-AUG-1998; 98US-0097955P.
PR 26-AUG-1998; 98US-0097971P.
PR 26-AUG-1998; 98US-0097974P.
PR 26-AUG-1998; 98US-0098014P.
PR 01-SEP-1998; 98US-0098716P.
PR 01-SEP-1998; 98US-0098723P.
PR 02-SEP-1998; 98US-0098803P.
PR 02-SEP-1998; 98US-0098821P.
PR 02-SEP-1998; 98US-0098843P.
PR 09-SEP-1998; 98US-0099602P.
PR 10-SEP-1998; 98US-0099741P.
PR 10-SEP-1998; 98US-0099754P.
PR 10-SEP-1998; 98US-0099763P.
PR 10-SEP-1998; 98US-0099812P.
PR 15-SEP-1998; 98US-0100388P.
PR 16-SEP-1998; 98US-0100662P.
PR 16-SEP-1998; 98US-0100664P.
PR 16-SEP-1998; 98US-0101751P.
PR 16-SEP-1998; 98WO-US019330.
PR 17-SEP-1998; 98US-0100683P.
PR 17-SEP-1998; 98US-0100684P.
PR 17-SEP-1998; 98US-0100919P.
PR 17-SEP-1998; 98US-0100930P.
PR 18-SEP-1998; 98US-0100849P.
PR 18-SEP-1998; 98US-0101014P.
PR 18-SEP-1998; 98US-0101068P.
PR 23-SEP-1998; 98US-0101471P.
PR 23-SEP-1998; 98US-0101472P.
PR 23-SEP-1998; 98US-0101475P.
PR 23-SEP-1998; 98US-0101477P.
PR 24-SEP-1998; 98US-0101738P.
PR 24-SEP-1998; 98US-0101739P.
PR 24-SEP-1998; 98US-0101743P.
PR 24-SEP-1998; 98US-0101922P.
PR 25-SEP-1998; 98US-0101786P.
PR 29-SEP-1998; 98US-0102207P.
PR 29-SEP-1998; 98US-0102240P.
PR 29-SEP-1998; 98US-0102330P.
PR 29-SEP-1998; 98US-0102331P.
PR 30-SEP-1998; 98US-0102487P.
PR 30-SEP-1998; 98US-0102570P.
PR 30-SEP-1998; 98US-0102571P.
PR 01-OCT-1998; 98US-0102684P.
PR 01-OCT-1998; 98US-0102687P.

Query Match 99.3%; Score 1986; DB 6; Length 382;
Best Local Similarity 97.9%; Pred. No. 4.8e-181;
Matches 374; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

```
Qy          1 MRPGTALQAVLLAVLLVGLRAATGRLLS--------GQPVCRGGTQRPCYKVIYFHDTSR 52
              ||||||||||||||||||||||||||||        ||||||||||||||||||||||||
Db          1 MRPGTALQAVLLAVLLVGLRAATGRLLSASDLDLRGGQPVCRGGTQRPCYKVIYFHDTSR 60

Qy         53 RLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTA 112
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 RLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTA 120

Qy        113 CQDLYAWTDGSISQFRNWYVDEPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKN 172
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CQDLYAWTDGSISQFRNWYVDEPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKN 180

Qy        173 NFICKYSDEKPAVPSREAEGEETELTTPVLPEETQEEDAKKTFKESREAALNLAYILIPS 232
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 NFICKYSDEKPAVPSREAEGEETELTTPVLPEETQEEDAKKTFKESREAALNLAYILIPS 240

Qy        233 IPLLLLLVVTTVVCWVWICRKRKREQPDPSTKKQHTIWPSPHQGNSPDLEVYNVIRKQSE 292
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 IPLLLLLVVTTVVCWVWICRKRKREQPDPSTKKQHTIWPSPHQGNSPDLEVYNVIRKQSE 300

Qy        293 ADLAETRPDLKNISFRVCSGEATPDDMSCDYDNMAVNPSESGFVTLVSVESGFVTNDIYE 352
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ADLAETRPDLKNISFRVCSGEATPDDMSCDYDNMAVNPSESGFVTLVSVESGFVTNDIYE 360

Qy        353 FSPDQMGRSKESGWVENEIYGY 374
              ||||||||||||||||||||||
Db        361 FSPDQMGRSKESGWVENEIYGY 382
```

Search completed: September 9, 2004, 22:21:29
Job time : 64 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: September 9, 2004, 22:56:51 ; Search time 90.7224 Seconds
(without alignments)
721.109 Million cell updates/sec

Title: US-09-887-855-2_COPY_24_227
Perfect score: 1106
Sequence: 1 GRLLSGQPVCRGGTQRPCYK.....EEDAKKTFKESREAALNLAY 204

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1335176 seqs, 320689617 residues

Total number of hits satisfying chosen parameters: 1335176

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1106	100.0	206	9	US-09-887-855-5	Sequence 5, Appli
2	1106	100.0	374	9	US-09-887-855-2	Sequence 2, Appli
3	1106	100.0	374	14	US-10-149-819-15	Sequence 15, Appl
4	1106	100.0	374	15	US-10-094-749-2090	Sequence 2090, Ap
5	1106	100.0	374	15	US-10-264-237-2579	Sequence 2579, Ap
6	1106	100.0	374	16	US-10-648-593-148	Sequence 148, App
7	1106	100.0	387	12	US-10-296-115-1311	Sequence 1311, Ap
8	1102	99.6	374	15	US-10-094-749-2142	Sequence 2142, Ap
9	1096	99.1	374	12	US-10-351-334-166	Sequence 166, App
10	1092	98.7	382	9	US-09-909-320-137	Sequence 137, App
11	1092	98.7	382	9	US-09-909-088B-137	Sequence 137, App
12	1092	98.7	382	9	US-09-905-291A-137	Sequence 137, App
13	1092	98.7	382	9	US-09-902-853-137	Sequence 137, App
14	1092	98.7	382	9	US-09-907-824-137	Sequence 137, App
15	1092	98.7	382	9	US-09-907-841-137	Sequence 137, App
16	1092	98.7	382	10	US-09-904-011-137	Sequence 137, App
17	1092	98.7	382	10	US-09-906-742-137	Sequence 137, App
18	1092	98.7	382	10	US-09-906-838-137	Sequence 137, App
19	1092	98.7	382	10	US-09-907-613-137	Sequence 137, App
20	1092	98.7	382	10	US-09-907-942-137	Sequence 137, App
21	1092	98.7	382	10	US-09-904-859-137	Sequence 137, App
22	1092	98.7	382	10	US-09-909-204-137	Sequence 137, App
23	1092	98.7	382	10	US-09-904-820-137	Sequence 137, App
24	1092	98.7	382	10	US-09-904-786-137	Sequence 137, App
25	1092	98.7	382	10	US-09-906-646-137	Sequence 137, App
26	1092	98.7	382	10	US-09-906-700-137	Sequence 137, App
27	1092	98.7	382	10	US-09-903-786-137	Sequence 137, App
28	1092	98.7	382	10	US-09-902-903-137	Sequence 137, App
29	1092	98.7	382	10	US-09-903-749A-137	Sequence 137, App
30	1092	98.7	382	10	US-09-904-119-137	Sequence 137, App
31	1092	98.7	382	10	US-09-904-956-137	Sequence 137, App
32	1092	98.7	382	10	US-09-902-736-137	Sequence 137, App
33	1092	98.7	382	10	US-09-907-794-137	Sequence 137, App
34	1092	98.7	382	10	US-09-903-943-137	Sequence 137, App
35	1092	98.7	382	10	US-09-904-462-137	Sequence 137, App
36	1092	98.7	382	10	US-09-907-925-137	Sequence 137, App
37	1092	98.7	382	10	US-09-902-692-137	Sequence 137, App
38	1092	98.7	382	10	US-09-903-520-137	Sequence 137, App
39	1092	98.7	382	10	US-09-905-056-137	Sequence 137, App
40	1092	98.7	382	10	US-09-909-064-137	Sequence 137, App
41	1092	98.7	382	10	US-09-904-553-137	Sequence 137, App
42	1092	98.7	382	10	US-09-905-381-137	Sequence 137, App
43	1092	98.7	382	10	US-09-905-088-137	Sequence 137, App
44	1092	98.7	382	10	US-09-907-575-137	Sequence 137, App
45	1092	98.7	382	10	US-09-905-075-137	Sequence 137, App

ALIGNMENTS

RESULT 1
US-09-887-855-5
; Sequence 5, Application US/09887855
; Patent No. US20020058310A1
; GENERAL INFORMATION:
; APPLICANT: Immunex Corporation
; APPLICANT: Anderson, Dirk M
; TITLE OF INVENTION: LECTIN SS3939 DNA AND POLYPEPTIDES
; FILE REFERENCE: 2883-US
; CURRENT APPLICATION NUMBER: US/09/887,855
; CURRENT FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 206
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-887-855-5

Query Match 100.0%; Score 1106; DB 9; Length 206;
Best Local Similarity 100.0%; Pred. No. 1.3e-103;
Matches 204; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 GRLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIE 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          3 GRLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIE 62

Qy         61 KPIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEPSCGSEVCV 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         63 KPIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEPSCGSEVCV 122

Qy        121 VMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGEETELTTPVLP 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        123 VMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGEETELTTPVLP 182

Qy        181 EETQEEDAKKTFKESREAALNLAY 204
              ||||||||||||||||||||||||
Db        183 EETQEEDAKKTFKESREAALNLAY 206
```

```
RESULT 2
US-09-887-855-2
; Sequence 2, Application US/09887855
; Patent No. US20020058310A1
; GENERAL INFORMATION:
;  APPLICANT: Immunex Corporation
;  APPLICANT:  Anderson, Dirk M
;  TITLE OF INVENTION: LECTIN SS3939 DNA AND POLYPEPTIDES
;  FILE REFERENCE: 2883-US
;  CURRENT APPLICATION NUMBER: US/09/887,855
;  CURRENT FILING DATE:  2001-06-22
;  NUMBER OF SEQ ID NOS: 9
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
;   LENGTH: 374
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-887-855-2

  Query Match           100.0%;  Score 1106;  DB 9;  Length 374;
  Best Local Similarity  100.0%;  Pred. No. 2.9e-103;
  Matches  204;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GRLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIE 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         24 GRLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIE 83

Qy         61 KFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEPSCGSEVCV 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         84 KFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEPSCGSEVCV 143

Qy        121 VMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGEETELTTPVLP 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        144 VMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGEETELTTPVLP 203

Qy        181 EETQEEDAKKTFKESREAALNLAY 204
              ||||||||||||||||||||||||
Db        204 EETQEEDAKKTFKESREAALNLAY 227


RESULT 3
US-10-149-819-15
; Sequence 15, Application US/10149819
; Publication No. US20030044913A1
; GENERAL INFORMATION:
;  APPLICANT: INCYTE GENOMICS, INC.
;  APPLICANT:  YUE, Henry
;  APPLICANT:  AZIMZAI, Yalda
;  APPLICANT:  TANG, Y. Tom
;  APPLICANT:  PATTERSON, Chandra
;  APPLICANT:  BAUGHN, Mariah R.
;  APPLICANT:  LU, Dyung Aina M.
;  APPLICANT:  SHAH, Purvi
;  APPLICANT:  LAL, Preeti
;  APPLICANT:  AU-YOUNG, Janice
;  APPLICANT:  BURFORD, Neil
;  TITLE OF INVENTION: EXTRACELLULAR MATRIX AND CELL ADHESION MOLECULES
;  FILE REFERENCE: PF-0760 PCT
;  CURRENT APPLICATION NUMBER: US/10/149,819
;  CURRENT FILING DATE:  2002-06-10
;  PRIOR APPLICATION NUMBER: 60/172,852; 60/172,354
;  PRIOR FILING DATE: 1999-12-10; 1999-12-16
;  NUMBER OF SEQ ID NOS: 42
;  SOFTWARE: PERL Program
; SEQ ID NO 15
;   LENGTH: 374
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No. US20030044913A1 3143411CD1
US-10-149-819-15

  Query Match           100.0%;  Score 1106;  DB 14;  Length 374;
  Best Local Similarity  100.0%;  Pred. No. 2.9e-103;
  Matches  204;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GRLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIE 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         24 GRLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIE 83

Qy         61 KFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEPSCGSEVCV 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         84 KFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEPSCGSEVCV 143

Qy        121 VMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGEETELTTPVLP 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        144 VMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGEETELTTPVLP 203

Qy        181 EETQEEDAKKTFKESREAALNLAY 204
              ||||||||||||||||||||||||
Db        204 EETQEEDAKKTFKESREAALNLAY 227


RESULT 4
US-10-094-749-2090
; Sequence 2090, Application US/10094749
; Publication No. US20030219741A1
; GENERAL INFORMATION:
;  APPLICANT: ISOGAI, TAKAO
;  APPLICANT:  SUGIYAMA, TOMOYASU
;  APPLICANT:  OTSUKI, TETSUJI
;  APPLICANT:  WAKAMATSU, AI
;  APPLICANT:  SATO, HIROYUKI
;  APPLICANT:  ISHII, SHIZUKO
;  APPLICANT:  YAMAMOTO, JUN-ICHI
;  APPLICANT:  ISONO, YUUKO
;  APPLICANT:  HIO, YURI
;  APPLICANT:  OTSUKA, KAORU
;  APPLICANT:  NAGAI, KEIICHI
;  APPLICANT:  IRIE, RYOTARO
;  APPLICANT:  TAMECHIKA, ICHIRO
;  APPLICANT:  SEKI, NAOHIKO
;  APPLICANT:  YOSHIKAWA, TSUTOMU
;  APPLICANT:  OTSUKA, MOTOYUKI
;  APPLICANT:  NAGAHARI, KENJI
;  APPLICANT:  MASUHO, YASUHIKO
;  TITLE OF INVENTION: NOVEL FULL-LENGTH cDNA
;  FILE REFERENCE: 084335/0160
;  CURRENT APPLICATION NUMBER: US/10/094,749
;  CURRENT FILING DATE:  2002-03-12
;  PRIOR APPLICATION NUMBER: 60/350,435
;  PRIOR FILING DATE: 2002-01-24
;  PRIOR APPLICATION NUMBER: JP 2001-328381
;  PRIOR FILING DATE: 2001-09-14
;  NUMBER OF SEQ ID NOS: 3381
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2090
;   LENGTH: 374
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-094-749-2090

  Query Match           100.0%;  Score 1106;  DB 15;  Length 374;
  Best Local Similarity  100.0%;  Pred. No. 2.9e-103;
  Matches  204;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GRLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIE 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         24 GRLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIE 83

Qy         61 KFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEPSCGSEVCV 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         84 KFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEPSCGSEVCV 143
```

```
Qy        121 VMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGEETELTTPVLP 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        144 VMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGEETELTTPVLP 203

Qy        181 EETQEEDAKKTFKESREAALNLAY 204
              ||||||||||||||||||||||||
Db        204 EETQEEDAKKTFKESREAALNLAY 227
```

RESULT 5
US-10-264-237-2579
; Sequence 2579, Application US/10264237
; Publication No. US20040009491A1
; GENERAL INFORMATION:
; APPLICANT: Birse et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA131P1
; CURRENT APPLICATION NUMBER: US/10/264,237
; CURRENT FILING DATE: 2002-10-04
; PRIOR APPLICATION NUMBER: PCT/US01/16450
; PRIOR FILING DATE: 2001-05-18
; PRIOR APPLICATION NUMBER: US 60/205,515
; PRIOR FILING DATE: 2000-05-19
; NUMBER OF SEQ ID NOS: 2876
; SOFTWARE: PatentIn Ver. 3.1
; SEQ ID NO 2579
; LENGTH: 374
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-264-237-2579

Query Match 100.0%; Score 1106; DB 15; Length 374;
Best Local Similarity 100.0%; Pred. No. 2.9e-103;
Matches 204; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 GRLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIE 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         24 GRLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIE 83

Qy         61 KFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEPSCGSEVCV 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         84 KFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEPSCGSEVCV 143

Qy        121 VMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGEETELTTPVLP 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        144 VMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGEETELTTPVLP 203

Qy        181 EETQEEDAKKTFKESREAALNLAY 204
              ||||||||||||||||||||||||
Db        204 EETQEEDAKKTFKESREAALNLAY 227
```

RESULT 6
US-10-648-593-148
; Sequence 148, Application US/10648593
; Publication No. US20040106132A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: IDENTIFICATION OF GENES FOR PREDICTING ACTIVITY OF COMPOUNDS THAT
; TITLE OF INVENTION: INTERACT WITH AND/OR MODULATE PROTEIN TYROSINE KINASES AND/OR
; TITLE OF INVENTION: PROTEIN TYROSINE KINASE PATHWAYS IN BREAST CELLS
; FILE REFERENCE: D0273 NP
; CURRENT APPLICATION NUMBER: US/10/648,593
; CURRENT FILING DATE: 2003-08-26
; PRIOR APPLICATION NUMBER: 60/406,385
; PRIOR FILING DATE: 2002-08-27
; NUMBER OF SEQ ID NOS: 557
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 148
; LENGTH: 374
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-648-593-148

Query Match 100.0%; Score 1106; DB 16; Length 374;
Best Local Similarity 100.0%; Pred. No. 2.9e-103;
Matches 204; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 GRLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIE 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         24 GRLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIE 83

Qy         61 KFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEPSCGSEVCV 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         84 KFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEPSCGSEVCV 143

Qy        121 VMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGEETELTTPVLP 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        144 VMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGEETELTTPVLP 203

Qy        181 EETQEEDAKKTFKESREAALNLAY 204
              ||||||||||||||||||||||||
Db        204 EETQEEDAKKTFKESREAALNLAY 227
```

RESULT 7
US-10-296-115-1311
; Sequence 1311, Application US/10296115
; Publication No. US20040053248A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq Inc
; TITLE OF INVENTION: No. US20040053248A1el Nucleic Acids and Polypeptides
; FILE REFERENCE: 784PCT
; CURRENT APPLICATION NUMBER: US/10/296,115
; CURRENT FILING DATE: 2002-11-18
; PRIOR APPLICATION NUMBER: US09/488,725
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: US09/552,317
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 1478
; SEQ ID NO 1311
; LENGTH: 387
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-296-115-1311

Query Match 100.0%; Score 1106; DB 12; Length 387;
Best Local Similarity 100.0%; Pred. No. 3e-103;
Matches 204; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 GRLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIE 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         37 GRLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIE 96

Qy         61 KFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEPSCGSEVCV 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         97 KFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEPSCGSEVCV 156

Qy        121 VMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGEETELTTPVLP 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        157 VMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGEETELTTPVLP 216

Qy        181 EETQEEDAKKTFKESREAALNLAY 204
              ||||||||||||||||||||||||
Db        217 EETQEEDAKKTFKESREAALNLAY 240
```

RESULT 8
US-10-094-749-2142
; Sequence 2142, Application US/10094749
; Publication No. US20030219741A1
; GENERAL INFORMATION:
; APPLICANT: ISOGAI, TAKAO
; APPLICANT: SUGIYAMA, TOMOYASU
; APPLICANT: OTSUKI, TETSUJI

; APPLICANT: WAKAMATSU, AI
; APPLICANT: SATO, HIROYUKI
; APPLICANT: ISHII, SHIZUKO
; APPLICANT: YAMAMOTO, JUN-ICHI
; APPLICANT: ISONO, YUUKO
; APPLICANT: HIO, YURI
; APPLICANT: OTSUKA, KAORU
; APPLICANT: NAGAI, KEIICHI
; APPLICANT: IRIE, RYOTARO
; APPLICANT: TAMECHIKA, ICHIRO
; APPLICANT: SEKI, NAOHIKO
; APPLICANT: YOSHIKAWA, TSUTOMU
; APPLICANT: OTSUKA, MOTOYUKI
; APPLICANT: NAGAHARI, KENJI
; APPLICANT: MASUHO, YASUHIKO
; TITLE OF INVENTION: NOVEL FULL-LENGTH cDNA
; FILE REFERENCE: 084335/0160
; CURRENT APPLICATION NUMBER: US/10/094,749
; CURRENT FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 60/350,435
; PRIOR FILING DATE: 2002-01-24
; PRIOR APPLICATION NUMBER: JP 2001-328381
; PRIOR FILING DATE: 2001-09-14
; NUMBER OF SEQ ID NOS: 3381
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2142
; LENGTH: 374
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-094-749-2142

Query Match 99.6%; Score 1102; DB 15; Length 374;
Best Local Similarity 99.5%; Pred. No. 7.4e-103;
Matches 203; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy       1 GRLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIE 60
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      24 GRLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIE 83

Qy      61 KFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEPSCGSEVCV 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      84 KFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEPSCGSEVCV 143

Qy     121 VMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGEETELTTPVLP 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     144 VMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGEETELTTPVLP 203

Qy     181 EETQEEDAKKTFKESREAALNLAY 204
           ||||||| ||||||||||||||||
Db     204 EETQEEDTKKTFKESREAALNLAY 227
```

RESULT 9
US-10-351-334-166
; Sequence 166, Application US/10351334
; Publication No. US20040034196A1
; GENERAL INFORMATION:
; APPLICANT: Komatsoulis et al
; TITLE OF INVENTION: 98 Human Secreted Proteins
; FILE REFERENCE: PZ031P2
; CURRENT APPLICATION NUMBER: US/10/351,334
; CURRENT FILING DATE: 2003-01-27
; PRIOR APPLICATION NUMBER: 60/350,898
; PRIOR FILING DATE: 2002-01-25
; PRIOR APPLICATION NUMBER: 09/489,847
; PRIOR FILING DATE: 2000-01-24
; PRIOR APPLICATION NUMBER: PCT/US99/17130
; PRIOR FILING DATE: 1999-07-29
; PRIOR APPLICATION NUMBER: 60/094,657
; PRIOR FILING DATE: 1998-07-30
; PRIOR APPLICATION NUMBER: 60/095,486
; PRIOR FILING DATE: 1998-08-05
; PRIOR APPLICATION NUMBER: 60/096,319
; PRIOR FILING DATE: 1998-08-12
; PRIOR APPLICATION NUMBER: 60/095,454
; PRIOR FILING DATE: 1998-08-06
; PRIOR APPLICATION NUMBER: 60/095,455
; PRIOR FILING DATE: 1998-08-06
; NUMBER OF SEQ ID NOS: 376
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 166
; LENGTH: 374
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (84)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (112)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-351-334-166

Query Match 99.1%; Score 1096; DB 12; Length 374;
Best Local Similarity 99.0%; Pred. No. 3e-102;
Matches 202; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy       1 GRLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIE 60
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      24 GRLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKLIE 83

Qy      61 KFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEPSCGSEVCV 120
           |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db      84 XFIENLLPSDGDFWIGLRRREEKQSNSTXCQDLYAWTDGSISQFRNWYVDEPSCGSEVCV 143

Qy     121 VMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGEETELTTPVLP 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     144 VMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGEETELTTPVLP 203

Qy     181 EETQEEDAKKTFKESREAALNLAY 204
           ||||||||||||||||||||||||
Db     204 EETQEEDAKKTFKESREAALNLAY 227
```

RESULT 10
US-09-909-320-137
; Sequence 137, Application US/09909320
; Patent No. US20020132240A1
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic

; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/909,320
; CURRENT FILING DATE: 2002-01-04
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 137
; LENGTH: 382
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-909-320-137

Query Match 98.7%; Score 1092; DB 9; Length 382;
Best Local Similarity 96.2%; Pred. No. 7.8e-102;
Matches 204; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

```
Qy          1 GRLLS--------GQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIES 52
              |||||        ||||||||||||||||||||||||||||||||||||||||||||||
Db         24 GRLLSASDLDLRGGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIES 83

Qy         53 EDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEP 112
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         84 EDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEP 143

Qy        113 SCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGEET 172
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        144 SCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGEET 203

Qy        173 ELTTPVLPEETQEEDAKKTFKESREAALNLAY 204
              ||||||||||||||||||||||||||||||||
Db        204 ELTTPVLPEETQEEDAKKTFKESREAALNLAY 235
```

RESULT 11
US-09-909-088B-137
; Sequence 137, Application US/09909088B
; Patent No. US20020146709A1
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/909,088B
; CURRENT FILING DATE: 2001-07-18
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 137
; LENGTH: 382
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-909-088B-137

Query Match 98.7%; Score 1092; DB 9; Length 382;
Best Local Similarity 96.2%; Pred. No. 7.8e-102;
Matches 204; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

```
Qy          1 GRLLS--------GQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIES 52
              |||||        ||||||||||||||||||||||||||||||||||||||||||||||||
Db         24 GRLLSASDLDLRGGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIES 83

Qy         53 EDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEP 112
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         84 EDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEP 143

Qy        113 SCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGEET 172
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        144 SCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGEET 203

Qy        173 ELTTPVLPEETQEEDAKKTFKESREAALNLAY 204
              ||||||||||||||||||||||||||||||||
Db        204 ELTTPVLPEETQEEDAKKTFKESREAALNLAY 235
```

RESULT 12
US-09-905-291A-137
; Sequence 137, Application US/09905291A
; Patent No. US20020160374A1
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/905,291A
; CURRENT FILING DATE: 2001-07-12
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 137
; LENGTH: 382
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-905-291A-137

Query Match 98.7%; Score 1092; DB 9; Length 382;
Best Local Similarity 96.2%; Pred. No. 7.8e-102;
Matches 204; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

```
Qy          1 GRLLS--------GQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIES 52
              |||||        ||||||||||||||||||||||||||||||||||||||||||||||||
Db         24 GRLLSASDLDLRGGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIES 83

Qy         53 EDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEP 112
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         84 EDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEP 143

Qy        113 SCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGEET 172
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        144 SCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGEET 203

Qy        173 ELTTPVLPEETQEEDAKKTFKESREAALNLAY 204
              ||||||||||||||||||||||||||||||||
Db        204 ELTTPVLPEETQEEDAKKTFKESREAALNLAY 235
```

RESULT 13
US-09-902-853-137
; Sequence 137, Application US/09902853
; Publication No. US20020192659A1
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same

```
;  FILE REFERENCE: 10466-14
;  CURRENT APPLICATION NUMBER: US/09/902,853
;  CURRENT FILING DATE:  2001-07-10
;  PRIOR APPLICATION NUMBER: US/09/665,350
;  PRIOR FILING DATE: 2000-09-18
;  PRIOR APPLICATION NUMBER: US 60/143,048
;  PRIOR FILING DATE: 1999-07-07
;  PRIOR APPLICATION NUMBER: US 60/145,698
;  PRIOR FILING DATE: 1999-07-26
;  PRIOR APPLICATION NUMBER: US 60/146,222
;  PRIOR FILING DATE: 1999-07-28
;  PRIOR APPLICATION NUMBER: PCT/US99/20594
;  PRIOR FILING DATE: 1999-09-08
;  PRIOR APPLICATION NUMBER: PCT/US99/20944
;  PRIOR FILING DATE: 1999-09-13
;  PRIOR APPLICATION NUMBER: PCT/US99/21090
;  PRIOR FILING DATE: 1999-09-15
;  PRIOR APPLICATION NUMBER: PCT/US99/21547
;  PRIOR FILING DATE: 1999-09-15
;  PRIOR APPLICATION NUMBER: PCT/US99/23089
;  PRIOR FILING DATE: 1999-10-05
;  PRIOR APPLICATION NUMBER: PCT/US99/28214
;  PRIOR FILING DATE: 1999-11-29
;  PRIOR APPLICATION NUMBER: PCT/US99/28313
;  PRIOR FILING DATE: 1999-11-30
;  PRIOR APPLICATION NUMBER: PCT/US99/28564
;  PRIOR FILING DATE: 1999-12-02
;  PRIOR APPLICATION NUMBER: PCT/US99/28565
;  PRIOR FILING DATE: 1999-12-02
;  PRIOR APPLICATION NUMBER: PCT/US99/30095
;  PRIOR FILING DATE: 1999-12-16
;  PRIOR APPLICATION NUMBER: PCT/US99/30911
;  PRIOR FILING DATE: 1999-12-20
;  PRIOR APPLICATION NUMBER: PCT/US99/30999
;  PRIOR FILING DATE: 1999-12-20
;  PRIOR APPLICATION NUMBER: PCT/US00/00219
;  PRIOR FILING DATE: 2000-01-05
;  NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 137
;    LENGTH: 382
;    TYPE: PRT
;    ORGANISM: Homo Sapien
US-09-902-853-137

  Query Match             98.7%;  Score 1092;  DB 9;  Length 382;
  Best Local Similarity   96.2%;  Pred. No. 7.8e-102;
  Matches  204;  Conservative    0;  Mismatches    0;  Indels    8;  Gaps    1;

Qy          1 GRLLS--------GQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIES 52
              |||||        ||||||||||||||||||||||||||||||||||||||||||||||
Db         24 GRLLSASDLDLRGGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIES 83

Qy         53 EDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEP 112
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         84 EDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEP 143

Qy        113 SCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGEET 172
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        144 SCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGEET 203

Qy        173 ELTTPVLPEETQEEDAKKTFKESREAALNLAY 204
              ||||||||||||||||||||||||||||||||
Db        204 ELTTPVLPEETQEEDAKKTFKESREAALNLAY 235


RESULT 14
US-09-907-824-137
; Sequence 137, Application US/09907824
; Publication No. US20020197671A1
; GENERAL INFORMATION:
;  APPLICANT: Genentech, Inc.
;  APPLICANT:  Ashkenazi, Avi
;  APPLICANT:  Botstein, David
;  APPLICANT:  Desnoyers, Luc
;  APPLICANT:  Eaton, Dan L.
;  APPLICANT:  Ferrara, Napoleone
;  APPLICANT:  Filvaroff, Ellen
;  APPLICANT:  Fong, Sherman
;  APPLICANT:  Gao, Wei-Qiang
;  APPLICANT:  Gerber, Hanspeter
;  APPLICANT:  Gerritsen, Mary E.
;  APPLICANT:  Goddard, A.
;  APPLICANT:  Godowski, Paul J.
;  APPLICANT:  Grimaldi, Christopher J.
;  APPLICANT:  Gurney, Austin L.
;  APPLICANT:  Hillan, Kenneth, J.
;  APPLICANT:  Kljavin, Ivar J.
;  APPLICANT:  Mather, Jennie P.
;  APPLICANT:  Pan, James
;  APPLICANT:  Paoni, Nicholas F.
;  APPLICANT:  Roy, Margaret Ann
;  APPLICANT:  Stewart, Timothy A.
;  APPLICANT:  Tumas, Daniel
;  APPLICANT:  Williams, P. Mickey
;  APPLICANT:  Wood, William, I.
;  TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
;  TITLE OF INVENTION:  Acids Encoding the Same
;  FILE REFERENCE: 10466-14
;  CURRENT APPLICATION NUMBER: US/09/907,824
;  CURRENT FILING DATE:  2001-07-17
;  PRIOR APPLICATION NUMBER: 09/665,350
;  PRIOR FILING DATE: 2000-09-18
;  PRIOR APPLICATION NUMBER: PCT/US00/04414
;  PRIOR FILING DATE: 2000-02-22
;  PRIOR APPLICATION NUMBER: US 60/143,048
;  PRIOR FILING DATE: 1999-07-07
;  PRIOR APPLICATION NUMBER: US 60/145,698
;  PRIOR FILING DATE: 1999-07-26
;  PRIOR APPLICATION NUMBER: US 60/146,222
;  PRIOR FILING DATE: 1999-07-28
;  PRIOR APPLICATION NUMBER: PCT/US99/20594
;  PRIOR FILING DATE: 1999-09-08
;  PRIOR APPLICATION NUMBER: PCT/US99/20944
;  PRIOR FILING DATE: 1999-09-13
;  PRIOR APPLICATION NUMBER: PCT/US99/21090
;  PRIOR FILING DATE: 1999-09-15
;  PRIOR APPLICATION NUMBER: PCT/US99/21547
;  PRIOR FILING DATE: 1999-09-15
;  PRIOR APPLICATION NUMBER: PCT/US99/23089
;  PRIOR FILING DATE: 1999-10-05
;  PRIOR APPLICATION NUMBER: PCT/US99/28214
;  PRIOR FILING DATE: 1999-11-29
;  PRIOR APPLICATION NUMBER: PCT/US99/28313
;  PRIOR FILING DATE: 1999-11-30
;  PRIOR APPLICATION NUMBER: PCT/US99/28564
;  PRIOR FILING DATE: 1999-12-02
;  PRIOR APPLICATION NUMBER: PCT/US99/28565
;  PRIOR FILING DATE: 1999-12-02
;  PRIOR APPLICATION NUMBER: PCT/US99/30095
;  PRIOR FILING DATE: 1999-12-16
;  PRIOR APPLICATION NUMBER: PCT/US99/30911
;  PRIOR FILING DATE: 1999-12-20
;  PRIOR APPLICATION NUMBER: PCT/US99/30999
;  PRIOR FILING DATE: 1999-12-20
;  PRIOR APPLICATION NUMBER: PCT/US00/00219
;  PRIOR FILING DATE: 2000-01-05
;  NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 137
;    LENGTH: 382
;    TYPE: PRT
;    ORGANISM: Homo Sapien
US-09-907-824-137

  Query Match             98.7%;  Score 1092;  DB 9;  Length 382;
  Best Local Similarity   96.2%;  Pred. No. 7.8e-102;
  Matches  204;  Conservative    0;  Mismatches    0;  Indels    8;  Gaps    1;
```

```
Qy          1 GRLLS--------GQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIES 52
              |||||        ||||||||||||||||||||||||||||||||||||||||||||||||
Db         24 GRLLSASDLDLRGGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIES 83

Qy         53 EDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEP 112
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         84 EDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEP 143

Qy        113 SCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGEET 172
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        144 SCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGEET 203

Qy        173 ELTTPVLPEETQEEDAKKTFKESREAALNLAY 204
              ||||||||||||||||||||||||||||||||
Db        204 ELTTPVLPEETQEEDAKKTFKESREAALNLAY 235


RESULT 15
US-09-907-841-137
; Sequence 137, Application US/09907841
; Publication No. US20020198366A1
; GENERAL INFORMATION:
;  APPLICANT: Genentech, Inc.
;  APPLICANT:  Ashkenazi, Avi
;  APPLICANT:  Botstein, David
;  APPLICANT:  Desnoyers, Luc
;  APPLICANT:  Eaton, Dan L.
;  APPLICANT:  Ferrara, Napoleone
;  APPLICANT:  Filvaroff, Ellen
;  APPLICANT:  Fong, Sherman
;  APPLICANT:  Gao, Wei-Qiang
;  APPLICANT:  Gerber, Hanspeter
;  APPLICANT:  Gerritsen, Mary E.
;  APPLICANT:  Goddard, A.
;  APPLICANT:  Godowski, Paul J.
;  APPLICANT:  Grimaldi, Christopher J.
;  APPLICANT:  Gurney, Austin L.
;  APPLICANT:  Hillan, Kenneth, J.
;  APPLICANT:  Kljavin, Ivar J.
;  APPLICANT:  Mather, Jennie P.
;  APPLICANT:  Pan, James
;  APPLICANT:  Paoni, Nicholas F.
;  APPLICANT:  Roy, Margaret Ann
;  APPLICANT:  Stewart, Timothy A.
;  APPLICANT:  Tumas, Daniel
;  APPLICANT:  Williams, P. Mickey
;  APPLICANT:  Wood, William, I.
;  TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
;  TITLE OF INVENTION:  Acids Encoding the Same
;  FILE REFERENCE: 10466-14
;  CURRENT APPLICATION NUMBER: US/09/907,841
;  CURRENT FILING DATE:  2001-11-20
;  PRIOR APPLICATION NUMBER: PCT/US00/04414
;  PRIOR FILING DATE: 2000-02-22
;  PRIOR APPLICATION NUMBER: US 60/143,048
;  PRIOR FILING DATE: 1999-07-07
;  PRIOR APPLICATION NUMBER: US 60/145,698
;  PRIOR FILING DATE: 1999-07-26
;  PRIOR APPLICATION NUMBER: US 60/146,222
;  PRIOR FILING DATE: 1999-07-28
;  PRIOR APPLICATION NUMBER: PCT/US99/20594
;  PRIOR FILING DATE: 1999-09-08
;  PRIOR APPLICATION NUMBER: PCT/US99/20944
;  PRIOR FILING DATE: 1999-09-13
;  PRIOR APPLICATION NUMBER: PCT/US99/21090
;  PRIOR FILING DATE: 1999-09-15
;  PRIOR APPLICATION NUMBER: PCT/US99/21547
;  PRIOR FILING DATE: 1999-09-15
;  PRIOR APPLICATION NUMBER: PCT/US99/23089
;  PRIOR FILING DATE: 1999-10-05
;  PRIOR APPLICATION NUMBER: PCT/US99/28214
;  PRIOR FILING DATE: 1999-11-29
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 137
;   LENGTH: 382
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-907-841-137

  Query Match             98.7%;  Score 1092;  DB 9;  Length 382;
  Best Local Similarity   96.2%;  Pred. No. 7.8e-102;
  Matches  204;  Conservative    0;  Mismatches    0;  Indels    8;  Gaps    1;

Qy          1 GRLLS--------GQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIES 52
              |||||        ||||||||||||||||||||||||||||||||||||||||||||||||
Db         24 GRLLSASDLDLRGGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIES 83

Qy         53 EDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEP 112
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         84 EDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEP 143

Qy        113 SCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGEET 172
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        144 SCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGEET 203

Qy        173 ELTTPVLPEETQEEDAKKTFKESREAALNLAY 204
              ||||||||||||||||||||||||||||||||
Db        204 ELTTPVLPEETQEEDAKKTFKESREAALNLAY 235


RESULT 16
US-09-904-011-137
; Sequence 137, Application US/09904011
; Publication No. US20030003530A1
; GENERAL INFORMATION:
;  APPLICANT: Genentech, Inc.
;  APPLICANT:  Ashkenazi, Avi
;  APPLICANT:  Botstein, David
;  APPLICANT:  Desnoyers, Luc
;  APPLICANT:  Eaton, Dan L.
;  APPLICANT:  Ferrara, Napoleone
;  APPLICANT:  Filvaroff, Ellen
;  APPLICANT:  Fong, Sherman
;  APPLICANT:  Gao, Wei-Qiang
;  APPLICANT:  Gerber, Hanspeter
;  APPLICANT:  Gerritsen, Mary E.
;  APPLICANT:  Goddard, A.
;  APPLICANT:  Godowski, Paul J.
;  APPLICANT:  Grimaldi, Christopher J.
;  APPLICANT:  Gurney, Austin L.
;  APPLICANT:  Hillan, Kenneth, J.
;  APPLICANT:  Kljavin, Ivar J.
;  APPLICANT:  Mather, Jennie P.
;  APPLICANT:  Pan, James
;  APPLICANT:  Paoni, Nicholas F.
;  APPLICANT:  Roy, Margaret Ann
;  APPLICANT:  Stewart, Timothy A.
;  APPLICANT:  Tumas, Daniel
;  APPLICANT:  Williams, P. Mickey
;  APPLICANT:  Wood, William, I.
;  TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
;  TITLE OF INVENTION:  Acids Encoding the Same
;  FILE REFERENCE: 10466-14
;  CURRENT APPLICATION NUMBER: US/09/904,011
;  CURRENT FILING DATE:  2001-07-11
;  PRIOR APPLICATION NUMBER: 09/665,350
;  PRIOR FILING DATE: 2000-09-18
;  PRIOR APPLICATION NUMBER: PCT/US00/04414
;  PRIOR FILING DATE: 2000-02-22
;  PRIOR APPLICATION NUMBER: US 60/143,048
;  PRIOR FILING DATE: 1999-07-07
;  PRIOR APPLICATION NUMBER: US 60/145,698
;  PRIOR FILING DATE: 1999-07-26
;  PRIOR APPLICATION NUMBER: US 60/146,222
```

; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 137
; LENGTH: 382
; TYPE: PRT
; ORGANISM: Homo Sapien
US-09-904-011-137

Query Match 98.7%; Score 1092; DB 10; Length 382;
Best Local Similarity 96.2%; Pred. No. 7.8e-102;
Matches 204; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

```
Qy          1 GRLLS--------GQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIES 52
              |||||        ||||||||||||||||||||||||||||||||||||||||||||||
Db         24 GRLLSASDLDLRGGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIES 83

Qy         53 EDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEP 112
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         84 EDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEP 143

Qy        113 SCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGEET 172
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        144 SCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGEET 203

Qy        173 ELTTPVLPEETQEEDAKKTFKESREAALNLAY 204
              ||||||||||||||||||||||||||||||||
Db        204 ELTTPVLPEETQEEDAKKTFKESREAALNLAY 235
```

RESULT 17
US-09-906-742-137
; Sequence 137, Application US/09906742
; Publication No. US20030023054A1
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/906,742
; CURRENT FILING DATE: 2001-07-16
; PRIOR APPLICATION NUMBER: 09/665,350
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 137
; LENGTH: 382
; TYPE: PRT
; ORGANISM: Homo Sapien
US-09-906-742-137

Query Match 98.7%; Score 1092; DB 10; Length 382;
Best Local Similarity 96.2%; Pred. No. 7.8e-102;
Matches 204; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

```
Qy          1 GRLLS--------GQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIES 52
              |||||        ||||||||||||||||||||||||||||||||||||||||||||||
Db         24 GRLLSASDLDLRGGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIES 83

Qy         53 EDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEP 112
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         84 EDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEP 143

Qy        113 SCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGEET 172
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        144 SCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGEET 203

Qy        173 ELTTPVLPEETQEEDAKKTFKESREAALNLAY 204
                ||||||||||||||||||||||||||||||||
Db        204 ELTTPVLPEETQEEDAKKTFKESREAALNLAY 235


RESULT 18
US-09-906-838-137
; Sequence 137, Application US/09906838
; Publication No. US20030027143A1
; GENERAL INFORMATION:
;  APPLICANT: Genentech, Inc.
;  APPLICANT:  Ashkenazi, Avi
;  APPLICANT:  Botstein, David
;  APPLICANT:  Desnoyers, Luc
;  APPLICANT:  Eaton, Dan L.
;  APPLICANT:  Ferrara, Napoleone
;  APPLICANT:  Filvaroff, Ellen
;  APPLICANT:  Fong, Sherman
;  APPLICANT:  Gao, Wei-Qiang
;  APPLICANT:  Gerber, Hanspeter
;  APPLICANT:  Gerritsen, Mary E.
;  APPLICANT:  Goddard, A.
;  APPLICANT:  Godowski, Paul J.
;  APPLICANT:  Grimaldi, Christopher J.
;  APPLICANT:  Gurney, Austin L.
;  APPLICANT:  Hillan, Kenneth, J.
;  APPLICANT:  Kljavin, Ivar J.
;  APPLICANT:  Mather, Jennie P.
;  APPLICANT:  Pan, James
;  APPLICANT:  Paoni, Nicholas F.
;  APPLICANT:  Roy, Margaret Ann
;  APPLICANT:  Stewart, Timothy A.
;  APPLICANT:  Tumas, Daniel
;  APPLICANT:  Williams, P. Mickey
;  APPLICANT:  Wood, William, I.
;  TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
;  TITLE OF INVENTION:  Acids Encoding the Same
;  FILE REFERENCE: 10466-14
;  CURRENT APPLICATION NUMBER: US/09/906,838
;  CURRENT FILING DATE:  2001-07-16
;  PRIOR APPLICATION NUMBER: 09/665,350
;  PRIOR FILING DATE: 2000-09-18
;  PRIOR APPLICATION NUMBER: PCT/US00/04414
;  PRIOR FILING DATE: 2000-02-22
;  PRIOR APPLICATION NUMBER: US 60/143,048
;  PRIOR FILING DATE: 1999-07-07
;  PRIOR APPLICATION NUMBER: US 60/145,698
;  PRIOR FILING DATE: 1999-07-26
;  PRIOR APPLICATION NUMBER: US 60/146,222
;  PRIOR FILING DATE: 1999-07-28
;  PRIOR APPLICATION NUMBER: PCT/US99/20594
;  PRIOR FILING DATE: 1999-09-08
;  PRIOR APPLICATION NUMBER: PCT/US99/20944
;  PRIOR FILING DATE: 1999-09-13
;  PRIOR APPLICATION NUMBER: PCT/US99/21090
;  PRIOR FILING DATE: 1999-09-15
;  PRIOR APPLICATION NUMBER: PCT/US99/21547
;  PRIOR FILING DATE: 1999-09-15
;  PRIOR APPLICATION NUMBER: PCT/US99/23089
;  PRIOR FILING DATE: 1999-10-05
;  PRIOR APPLICATION NUMBER: PCT/US99/28214
;  PRIOR FILING DATE: 1999-11-29
;  PRIOR APPLICATION NUMBER: PCT/US99/28313
;  PRIOR FILING DATE: 1999-11-30
;  PRIOR APPLICATION NUMBER: PCT/US99/28564
;  PRIOR FILING DATE: 1999-12-02
;  PRIOR APPLICATION NUMBER: PCT/US99/28565
;  PRIOR FILING DATE: 1999-12-02
;  PRIOR APPLICATION NUMBER: PCT/US99/30095
;  PRIOR FILING DATE: 1999-12-16
;  PRIOR APPLICATION NUMBER: PCT/US99/30911
;  PRIOR FILING DATE: 1999-12-20
;  PRIOR APPLICATION NUMBER: PCT/US99/30999
;  PRIOR FILING DATE: 1999-12-20
;  PRIOR APPLICATION NUMBER: PCT/US00/00219
;  PRIOR FILING DATE: 2000-01-05
;  NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 137
;   LENGTH: 382
;   TYPE: PRT
;   ORGANISM: Homo Sapien
US-09-906-838-137

  Query Match             98.7%;  Score 1092;  DB 10;  Length 382;
  Best Local Similarity   96.2%;  Pred. No. 7.8e-102;
  Matches  204;  Conservative    0;  Mismatches    0;  Indels    8;  Gaps    1;

Qy          1 GRLLS--------GQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIES 52
                |||||        ||||||||||||||||||||||||||||||||||||||||||||
Db         24 GRLLSASDLDLRGGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIES 83

Qy         53 EDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEP 112
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         84 EDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEP 143

Qy        113 SCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGEET 172
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        144 SCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGEET 203

Qy        173 ELTTPVLPEETQEEDAKKTFKESREAALNLAY 204
                ||||||||||||||||||||||||||||||||
Db        204 ELTTPVLPEETQEEDAKKTFKESREAALNLAY 235


RESULT 19
US-09-907-613-137
; Sequence 137, Application US/09907613
; Publication No. US20030027145A1
; GENERAL INFORMATION:
;  APPLICANT: Genentech, Inc.
;  APPLICANT:  Ashkenazi, Avi
;  APPLICANT:  Botstein, David
;  APPLICANT:  Desnoyers, Luc
;  APPLICANT:  Eaton, Dan L.
;  APPLICANT:  Ferrara, Napoleone
;  APPLICANT:  Filvaroff, Ellen
;  APPLICANT:  Fong, Sherman
;  APPLICANT:  Gao, Wei-Qiang
;  APPLICANT:  Gerber, Hanspeter
;  APPLICANT:  Gerritsen, Mary E.
;  APPLICANT:  Goddard, A.
;  APPLICANT:  Godowski, Paul J.
;  APPLICANT:  Grimaldi, Christopher J.
;  APPLICANT:  Gurney, Austin L.
;  APPLICANT:  Hillan, Kenneth, J.
;  APPLICANT:  Kljavin, Ivar J.
;  APPLICANT:  Mather, Jennie P.
;  APPLICANT:  Pan, James
;  APPLICANT:  Paoni, Nicholas F.
;  APPLICANT:  Roy, Margaret Ann
;  APPLICANT:  Stewart, Timothy A.
;  APPLICANT:  Tumas, Daniel
;  APPLICANT:  Williams, P. Mickey
;  APPLICANT:  Wood, William, I.
;  TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
;  TITLE OF INVENTION:  Acids Encoding the Same
;  FILE REFERENCE: 10466-14
;  CURRENT APPLICATION NUMBER: US/09/907,613
;  CURRENT FILING DATE:  2001-07-17
;  PRIOR APPLICATION NUMBER: PCT/US00/04414
;  PRIOR FILING DATE: 2000-02-22
;  PRIOR APPLICATION NUMBER: US 60/143,048
;  PRIOR FILING DATE: 1999-07-07
```